Investigating N-glycosylation sites in follicular lymphoma surface immunoglobulin and their role in lymphoma initiation

Mariette Odabashian

Submitted in partial fulfilment of the requirements of the Degree of Doctor of Philosophy

March 2019

Centre for Haemato-Oncology

Barts Cancer Institute, London, UK

Abstract

Follicular lymphoma (FL) is an incurable indolent non-Hodgkin's lymphoma that is both biologically and clinically heterogeneous. Despite the loss of one immunoglobulin allele through the t14;18 translocation and ongoing somatic hypermutation (SHM) of the immunoglobulin heavy chain variable region genes (*IGHV*), all detectable tumour subclones retain functional expression of the surface immunoglobulin throughout disease, resulting in thousands of tumour subclones displaying distinct but clonally related *IGHV* sequences. This retention suggests a tumour dependence on signalling through the B cell receptor (BCR).

>90% of FL cases contain N-glycosylation (N-gly) motifs in their *IGHV* which are acquired through SHM. Motifs are rarely found in normal somatically mutated B cells, indicative of a pathogenic function. Mannoses attach to motifs and activate BCR signalling pathways following engagement with lectins DC-SIGN and MR, expressed by microenvironmental immune cells. This novel interaction could represent a critical mechanism by which tumour cells survive in the hostile microenvironment.

Utilising a next-generation sequencing approach, we determined using temporal tumour samples from a diverse group of patients the behaviour of N-gly motifs during tumour evolution and progression to elucidate their importance in lymphomagenesis. Remarkably, we observed conservation of N-gly motif sites in >97% of subclone populations within and across disease events of all patients, inferring motifs are an early and stable event of pathogenesis. Rare motif negative subclones were presumably lost or negligible from successive events in contrast to motif positive subclones that were able to migrate between anatomical sites. This positive selection suggests an ongoing reliance of N-gly sites during disease progression, despite the changing genomic landscape of tumour clones.

Interrogation of the microenvironmental landscape through novel, multiplex phenotyping analysis provided new insight into DC-SIGN and MR expression on immune cells of healthy and FL derived lymph nodes. A greater population of follicular dendritic cells, macrophages and M2 macrophages within tumour lymph nodes expressed a MR+ and DCSIGN+MR+ phenotype, indicating a remodelling of the microenvironment in disease that could be contributed by the mannose- lectin interaction. Targeting motifs or the inferred mannose-lectin interaction may lead to therapeutic benefit.

Statement of originality

I, Mariette Odabashian, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

Signature: 

Date: 06/03/2019

Details of collaboration and publications:

Contents

Acknowledgements

My time at Barts Cancer Institute has been an extremely enjoyable experience, regarding both the work within this thesis and the wonderful people I have met along the way.

My immense gratitude goes to my supervisor Sergey Krysov for patiently and enthusiastically guiding me throughout the PhD, whose door was always open for discussion and support. It has been a pleasure working together and I am thankful to have had such a great mentor. My appreciation goes to my primary supervisor, Mariarita Calaminici, for her expert advice and continual encouragement, and for giving me the opportunity to pursue a doctorate.

Work in this thesis would not have been possible without the contribution from several people. Firstly, I would like to thank Jessica Okosun for access to patient samples. To Emanuela Carlotti, for the bioinformatics assistance and always offering help when I ran into issues with the sequencing data. To Andrew Clear, for patiently preparing numerous tissue sections and going through the patient databank to retrieve clinical information. To Joseph Taylor, for teaching me how to use the new immunohistochemistry digital analysis software and offering solutions to the numerous technical issues I ran into.

My PhD would not have been half as enjoyable without the friends I have made along the way. To Tahrima, Hemalvi, Faith and Emma, I look back fondly on our time in the office and the amount of laughter we have shared over the years. I am lucky to have made such lifelong friends. To Filomena Spada, it was a joy to work with such a positive person, thanks for all the helpful advice and discussions. To my family, who encouraged me to pursue a PhD and supported me through both the peaks and troughs of academia. To my partner Michael, for his understanding and faith in my abilities, and for offering words of encouragement when they were needed most.

Finally, I thank The Pathological Society of Great Britain and Ireland for generously supporting my PhD and for the patients that so openly support and contribute to our research endeavours.

Dedication

This thesis is dedicated to Mum & Dad who have always encouraged me to learn and take full advantage of the opportunities that were not available to them. Thank you for always supporting me in every decision I have made and having belief that I would eventually find my path. You are my biggest inspiration and motivation and I love you both very much.

Publications

Odabashian M, Carlotti E, Araf S, Okosun J, Forconi F, Stevenson FK, Gribben J, Calaminici M, Krysov S. Immunoglobulin Variable Region Gene Sequences Reveal N-Glycosylation Motifs As an Early and Stable Event in Follicular Lymphoma Pathology. https://ash.confex.com/ash/2018/webprogram/Paper112397.html

Araf S, Wang J, Ashton-Key M, Korfi K, Di Bella D, Rio-Machin A, **Odabashian M**, Foria V, Du MQ, Cucco F, Barrans S, Johnson P, Laird SR, Fisher AM, Cullis J, Graham TA, Okosun J, Fitzgibbon J, Chiecchio L. Transmission of diffuse large B cell lymphoma by an allogeneic stem-cell transplant. *Haematologica*. 2018.196907.

List of Figures

List of Tables

List of abbreviations

AID:	Activation induced cytidine deaminase
AKT:	Protein kinase B
BAFF:	B cell activating factor
BCR:	B cell receptor
BCL2:	B cell lymphoma-2
CCL17:	Chemokine (C-C motif) ligand 17
CDRs:	Complementarity determining regions
CLL:	Chronic lymphocytic leukaemia
CPC:	Common progenitor clone
CREBBP:	CREB binding protein
CSR:	Class switch recombination
C-type:	Calcium dependent
CNV:	Copy number variation
CXCL12/SDF-1:	C-X-C motif chemokine 12/ stromal cell-derived factor 1
CXCL13:	Chemokine (C-X-C motif) ligand 13
DC:	Dendritic cell
DC-SIGN:	Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin
DLBCL:	Diffuse large B cell lymphoma
DLI:	Donor leukocyte infusion
DNA:	Deoxyribonucleic acid
ER:	Endoplasmic reticulum
ERK:	Extracellular signal–regulated kinases
EZH2:	Enhancer of zeste homolog 2
FDC:	Follicular dendritic cell
FFPE:	Fresh frozen paraffin-embedded
FFS:	Five year failure free survival
FL:	Follicular lymphoma
FLLC:	Follicular lymphoma-like cells
GC:	Germinal centre
HH:	Heteroduplex analysis
HIER:	Heat induced epitope retrieval
HRP:	Horseradish peroxidase
IGH:	Immunoglobulin heavy chain gene
IGHV:	Immunoglobulin heavy chain variable gene
IL-4:	Interleukin 4
IL-6:	Interleukin 6
IL-10:	Interleukin 10
ISFN:	In situ follicular neoplasia
LN:	Lymph node
M2:	Alternatively activated macrophage
MBL:	Mannose binding lectin
MC:	Major clone

MHC:	Major histocompatibility complex
MR:	Mannose receptor
MSC:	Mesenchymal stromal cells
N-gly:	N-linked glycosylation
NHL:	Non-Hodgkin's Lymphoma
OS:	Overall survival
PFL:	Partial involvement by FL
PFS:	Progression free survival
PRR:	Pattern recognition receptors
PTFL:	Paediatric-type follicular lymphoma
POD:	Progression of disease
RAG:	Recombination activating gene
R-CHOP:	Rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone
SHM:	Somatic hypermutation
sIg:	Surface immunoglobulin
SNP:	Single nucleotide polymorphism
SNV:	Single nucleotide variant
STAT6	Signal transducer and activator of transcription 6
TAM:	Tumour associated macrophage
T_{FH}:	Follicular helper T cells
tFL:	Transformed follicular lymphoma
TNFRSF14:	Tumour necrosis factor receptor superfamily, member 14
T_{Regs}:	T regulatory cells
V:	Variable region gene
WES:	Whole exome sequencing
WGS:	Whole genome sequencing
W/W:	Watch and wait

Chapter 1: INTRODUCTION

1.1 Disease overview

Follicular lymphoma (FL) is the second most common non-Hodgkin's lymphoma (NHL) in Western Europe and the United States following diffuse large B cell lymphoma (DLBCL). The disease represents 70% of all indolent lymphomas with an increasing incidence in Western countries over the last two decades, rising from 2 in 100,000 people in the 1950s to 5 in 100,000 people in 2016 (Campo et al., 2011, Mounier et al., 2015). The majority of patients are diagnosed at an advanced stage due to lack of symptoms, with lymphadenopathy being the first reason for seeking medical assistance. Although there is no preponderance regarding sex, there are variations in incidence rates between ethnic groups and geographical location, the reason for this remaining unknown. The median age at diagnosis is 65 years (Junlen et al., 2015), highlighting FL as a disease strongly linked to age. Although rare, FL can manifest in young adults and children which is now recognised as a separate disease entity called paediatric-type FL (PTFL) due to being both histologically and clinically distinct (Pinto et al., 1990, Winberg et al., 1981, Lorsbach et al., 2002, Araf and Fitzgibbon, 2016).

The clinical course of FL is heterogeneous, with the majority of patients having a high response rate to first-line therapy followed by several cycles of relapse and remission with each disease free episode becoming progressively shorter. These patients usually follow an indolent course, surviving up to twenty years following diagnosis. However, a subset of patients (20-25%) follow an aggressive disease course in which progression occurs within two years of treatment or the disease transforms to a high-grade lymphoma (Casulo et al., 2015, Casulo, 2016). For these patients, the prognostic outlook is relatively very poor despite an availability of several immunotherapeutic strategies. Predicting which patients would fall into this high-risk group remains an ongoing area of research (Huet et al., 2018b) in order to deliver a tailored therapeutic strategy and enrol patients most likely to benefit from a clinical trial (Weigert and Weinstock, 2017).

Although we are beginning to unravel the biological processes underpinning the disease and targeting them with novel immunomodulatory therapeutic approaches, the disease effectively remains incurable.

1.2 Clinical presentation and diagnosis

The majority of patients present with lymphadenopathy in the cervical, axillary, inguinal and/or femoral regions. Rarely are tumours localised to one site and therefore widespread dissemination of disease is a common finding. Adenopathy waxes and wanes spontaneously. Apart from this enlargement, the majority of patients are asymptomatic. Less than 20% of patients display 'B symptoms' including night sweats, fever and/or weight loss. Despite the large tumour burden, increased serum lactate dehydrogenase are observed in less than 25% of patients. Extra-nodal involvement is common with liver and bone marrow involvement seen in 50% and 80% of cases, respectively.

A biopsy taken from the affected lymph node is preferably used in FL diagnosis, with core needle biopsy reserved for cases with difficult to access nodes (e.g. retroperitoneal nodes). Histologic analysis typically reveals a nodular growth pattern with abolished nodal architecture, including an ill-defined or absent mantle zone. FL is composed of follicular centre cells, named centrocytes and centroblasts. Grading, according to WHO classification, is based on the number of centroblasts counted per high-powered field (hpf), with an higher count indicative of a more aggressive clinical course (Table 1.1). Grade 3 is subdivided into group A and B (absence of centrocytes) with 3B resembling de novo DLBCL, both histologically and molecularly (Horn et al., 2011, Bosga-Bouwer et al., 2003). Because of this marked distinction, grade 3B is likely to be treated similarly to DLBCL.

Immunophenotyping of tissue sections is used to identify the disease. Typically, tumour cells express pan-B cell markers (IgM, CD19, CD20, CD22 AND CD79a) and germinal centre markers, including BCL-6 and CD10. They are negative for CD5 expression which is useful in differential diagnosis of low-grade B cell lymphomas (Cossman et al., 1984,

Harris et al., 1984, de Leon et al., 1998, Lydyard et al., 1999). Bcl-2 is overexpressed in the germinal centres of >90% of patients which is the result of the translocation t(14;18)(q32;q21) involving the *BCL2* gene and the immunoglobulin heavy chain enhancer, discussed in more detail below (Aster and Longtine, 2002). The surface immunoglobulin is monoclonal, with malignant cells sharing identical *V*, *D* and *J* genes in the variable region of both their heavy and light chains which are readily detected through the polymerase chain reaction (PCR). In accordance with their germinal centre residence, on-going somatic hypermutation is observed in the variable regions of malignant cells, giving rise to a heterogeneous clonal population (Cleary et al., 1986b, Bahler et al., 1991).

Grade	Definition
1	0-5 centroblasts/hpf
2	6-15 centroblasts/hpf
3A	>15 centroblasts/hpf Presence of centrocytes
3B	>15 centroblasts /hpf, sheets of centroblasts Absence of centrocytes

Table 1.1: *WHO classification of FL histological grades*. The system is based on the absolute number of centroblasts per 40x high powered field (hpf).

1.3 Staging and prognostic indicators

Anatomical staging is based on the Ann Arbour classification and describes the number of lymph nodes involved, their location and presence of extranodal organ involvement. Each stage is subdivided to indicate the absence or presence of B symptoms.

The clinical outcome for FL patients remains highly variable, with a subset of patients having a very poor clinical outcome, associated with histological transformation or progression of disease within 2 years of initial therapy (Casulo et al., 2015, Conconi et al., 2012). Identifying patients who are in this high-risk group can lead to tailoring of therapy and an improved clinical outcome, and therefore developing prognostic tools

that stratify patients based on disease-specific features is highly desirable. Follicular Lymphoma International Prognostic Index (FLIPI) predicts the outcome of patients through five independent adverse prognostic risk factors (9) (Table 1.2). Patients are placed into three risk groups, based on the number of risk factors they present. Because FLIPI was based on retrospective analysis before the introduction of widespread chemotherapy, FLIPI2 (10) was introduced which revised the adverse prognostic risk factors (Table 1.2) and is based on a progression-free survival (PFS) endpoint. However, the utility of the FLIPI-2 model in a clinical setting remains undetermined, with FLIPI continuing to be a useful prognostic model in the Rituximab era.

M7FLIPI is the first prognostic score in lymphoma to incorporate both genetic and clinical factors to identify high risk patients treated with standard therapy. It consists of the FLIPI risk factors, Eastern Cooperative Oncology Group performance status, and mutations in seven genes involved in epigenetic regulation (EZH2, EP300, and CREBBP), transcription (FOXO1, MEF2B), nucleosome remodelling (ARID1A), and BCR signalling (CARD11). The risk stratification is more closely associated with outcome compared to the clinical or genetic predictors alone.

Independent Risk factors	Risk group	No. of factors	Outcome, %	
			5-y OS	**10-y OS**
FLIPI				
Age ≥60 y	Low	0-1	91	71
Ann Arbor Stage III/IV				
Haemoglobin <12 g/dL	Intermediate	2	78	51
LDH elevated				
>4 nodal sites	High	3-5	53	36
			3-y PFS	**5-y PFS**
FLIPI-2				
Age ≥60 y	Low	0	91	80
B2M elevated				
Haemoglobin < 2 g/dL	Intermediate	1-2	69	51
BM involvement				
Lymph node diameter >6 cm	High	3-5	51	19

Table 1.2: *FLIPI and FLIPI-2 prognostic indexes*. FLIPI-2 was created in the rituximab therapy era and prognosis is based on progression-free survival rather than overall survival. FLIPI-2 also modified the risk factors of FLIPI, including beta 2 microglobulin (B2M) level, bone marrow (BM) involvement and lymph node diameter of >6cm.

1.4 Management strategies

Therapeutic choices are influenced by a range of factors including stage of disease, tumour burden, comorbidities, performance status, age and toxicity concerns. The Groupe d'Etude des Lymphomes Folliculaires (GELF) Criteria (Brice et al., 1997) can help identify patients who are at higher risk for rapid disease progression and therefore require immediate therapy.

1.4.1 Initial treatment of early stage FL

Less than 10% of patients are diagnosed with Stage I/II disease and due to this low number, no randomised studies regarding optimal management exist. However, targeted radiotherapy to the disease site has high curative potential and is generally the treatment of choice. 10 year overall survival rates based on this treatment modality alone is ~ 80% with a median survival of ~19 years (Guadagnolo et al., 2006). The recommended dose is currently 24 Gy with higher doses offering no benefit with regards to progression-free survival (PFS) or overall survival (OS) (Lowry et al., 2011).

1.4.2 Initial treatment of advanced stage FL

The majority of FL patients have advanced disease at diagnosis with either a low or high tumour burden, as assessed through the GELF criteria. For this group, the disease cannot be cured by conventional therapies, with the emphasis of treatment being to control symptoms and extend remission lengths and OS. Patients can be symptomatic or asymptomatic, which influences treatment decisions.

Patients with asymptomatic, low tumour burden generally undergo a watch and wait (W/W) approach. The patient is observed for signs of vital organ compromise and infiltration, development of cytopenias, B symptoms or rapid disease progression. Randomised phase III studies revealed that patients undergoing Rituximab induction followed by 2 years of rituximab maintenance had a significantly longer PFS when compared to the W/W group. However, there was no difference in OS between the two groups, maintaining W/W as a reasonable approach, which shields patients from the toxic effects of chemotherapy. However, Rituximab may be beneficial to a subset of patients who have coping issues and prefers a longer remission duration with an understanding of the efficacy/toxicity trade-off.

Patients with symptomatic, low tumour burden are first examined for alternative explanations for their symptoms. If no other cause is ascertained, single agent Rituximab

is used, particularly in older patients and those with co-morbidities. Rituximab based chemotherapy would likely benefit younger and fitter patients.

Patients with symptomatic, high tumour burden usually require immediate therapy. The inclusion of Rituximab to a variety of combination chemotherapy has been a considerable therapeutic advance in FL treatment as evidenced by improved response rates (RR), PFS, and OS (Hiddemann et al., 2005, Marcus et al., 2008, Herold et al., 2003). Rituximab is a chimeric anti-CD20 monoclonal antibody with proposed mechanisms of action including elicitation of antibody-dependent cellular cytotoxicity (ADCC) and complement mediated cytotoxicity following binding to CD20 on the B cell surface (Smith, 2003). Chemotherapy has a less targeted mechanism of action, with the focus being on interrupting DNA synthesis and cell division processes (Table 1.3). To evaluate which chemotherapy was optimal for first line treatment of advanced stage FL, R-CHOP (rituximab, cyclophosphamide, doxorubicin, vincristine, prednisolone) treatment was compared to R-CVP (rituximab, cyclophosphamide, vincristine, prednisolone) and R-FM (rituximab, fludarabine and mitoxantrone) (Federico et al., 2013). R-CHOP and R-FM had a better time to treatment failure (TTF) and PFS than R-CVP. However, R-FM was associated with a higher toxicity risk profile compared to R-CHOP and R-CVP. Yet, the three year OS rate for all categories was 95%, resulting in a lack of consensus on the choice of standard chemotherapy. In a German study group, R-CHOP was compared to bendamustine plus rituximab (BR) in a phase III trial in 513 advanced stage FL patients. BR had superior toxicity profiles and PFS at 45 months (Rummel et al., 2013). In the Bright study, BR was found to be non-inferior to R-CHOP or R-CVP in treatment naïve patients, measured by complete response rate (CRR) and overall response rate (ORR) (Flinn et al., 2014), establishing BR as a standard frontline therapy for indolent NHLs in North America and parts of Europe (Morschhauser et al., 2011, Rummel et al., 2013). The use of maintenance rituximab is individualised as although it improves the 2 year PFS, there are no benefits regarding OS (Salles et al., 2011).

Chemotherapy agent	Mechanism of action
Cyclophosphamide	*Alkylating agent*. Attachment of alkyl groups to DNA bases, preventing DNA synthesis
Doxorubicin	*Anthracycline*. Inhibits topoisomerase II activity
Vincristine	*Vinca alkaloid*. Binds to microtubular proteins of the mitotic spindle causing mitotic arrest at metaphase
Prednisolone	*Glucocorticoid*. Mediate pro-apoptotic pathways through transcriptional activation of the Bim protein
Fludarabine	Inhibits DNA synthesis by inhibiting DNA polymerase alpha, ribonucleotide reductase and DNA primase
Mitoxantrone	DNA intercalating agent causing strand breaks. Inhibitor of toposiomerase II
Bendamustine	*Alkylating agent.* Causes intra- and inter-strand crosslinks between DNA bases resulting in cell death

Table 1.3: *Chemotherapy agents used in combination with Rituximab.*

1.4.3 Treatment of relapsed FL

Despite the long remission and high response rates to first line treatment, the majority of patients inevitably relapse. To determine the extent of disease reoccurrence, a repeat biopsy is taken to determine the staging and exclusion of histological transformation. Considerations regarding therapy include previous treatments, duration of response, health and fitness of the patient and their current symptoms. There are a number of options for relapsed FL including radiotherapy for localised disease and for selected patients, allogeneic or autologous stem cell transplant. Rituximab with chemotherapy was shown to have an enhanced efficacy of treatment of relapsed FL not previously treated with rituximab or anthracycline. Patients were randomised to either CHOP or R-CHOP and those showing response either underwent rituximab maintenance or observation (van Oers et al., 2006). The R-CHOP group had an increased overall response rate and CR rate compared to CHOP alone with both groups benefiting with an improved PFS following rituximab maintenance compared to overall survival. Rituximab added to other chemotherapy such as fludarabine, cyclophosphamide, mitoxantrone (R-FCM) also showed a clinical benefit in a relapsed setting with superior PFS and OS compared

to FCM treatment alone (Forstpointner et al., 2004). Bendamustine as a single agent or in combination with rituximab also showed benefit to patients with relapsed FL (Friedberg et al., 2008, Robinson et al., 2008, Rummel et al., 2005).

Autologous stem cell transplant (ASCT) has been explored as part of consolidation therapy with high dose chemotherapy in high risk and advanced stage FL patients at first remission (Benedetti et al., 2013, Gyan et al., 2009, Ladetto et al., 2008, Lenz et al., 2004, Sebban et al., 2008) with an improved PFS. However, the lack of advantage regarding OS has not merited the use of HSCT as a first line therapy. In a relapsed setting however, high dose therapy with ASCT was shown to increase PFS and OS when compared to conventional chemotherapy (Schouten et al., 2003) in the pre-rituximab era. Prospective studies have indicated the promise of ASCT with rituximab as a treatment in the first relapse setting, as indicated by a 90% survival after relapse (SAR) at 5 years (Le Gouill et al., 2011, Sebban et al., 2008). Total body irradiation (TBI) as a conditioning regimen is not typically used due to a higher incidence of treatment related myelodysplastic syndrome/acute myeloid leukaemia compared to chemotherapy based conditioning (Montoto et al., 2007a, Darrington et al., 1994), outweighing the benefit of lowered relapsed risk.

1.4.4 Novel therapeutics

Targeted therapy exploiting the key mechanisms used by tumour cells has become a popular area of research and as a result, a wide range of novel agents are currently in development or in a clinical trial. Where available, relapsed/refractory patients requiring treatment should be entered into clinical trials evaluating novel therapies. A non-exhaustive summary of novel agents is provided in Table 1.4.

Agent	Classification	N	ORR, %	CR, %	1-y PFS, %
Lenalidomide	Immunomodulator	45	53	20	55
Lenalidomide + rituximab		46	76	39	80
Duvelisib (IPI-145)	PI3K-δ and PI3K-γ inhibitor	13	69	38	80
Ibrutinib	BTK inhibitor	40	28	5	50
Venetoclax (ABT-199)	BCL-2 inhibitor	29	38	14	40
Polatuzumab vedotin + rituximab	Anti-CD79b antibody-drug conjugate	45	73	33	63
Obinutuzumab	Anti-CD20 monoclonal antibody	74	45	12	65

Table 1.4: *Novel agents undergoing evaluation in relapsed FL patients with outcome data provided*. N (number of patients), ORR (overall response rate), CR (complete remission). Table information taken from Kahl & Yang, 2016 (Kahl and Yang, 2016).

1.5 FL variants

Whilst the majority of FL patients display a 'classical' nodal FL, less common variants occur which are both clinically and pathologically distinct. Below are examples of two FL variants.

Paediatric-type FL (PTFL): Typically presents in adolescents and young adults (18-30 years old), with an occasional presentation in older adults. There is a predominance in males (10:1) and the majority of patients present with localised stage one

lymphadenopathy. Prognosis is favourable following local excision or minimal chemotherapy. The disease has a follicular architecture composed of blastoid cells which are larger than centrocytes and centroblasts (Liu et al., 2013). Although having a high-grade histological appearance and high proliferation index, the behaviour is usually benign.

The genetic landscape of PTFL is also distinct from classical FL. Recurrent mutations in epigenetic modifiers including CREBBP and KMT2D are absent (Louissaint et al., 2016, Schmidt et al., 2016) in addition to a lack of BCL2, BCL6 and IRF4 rearrangements which are typical of classical FL (Louissaint et al., 2012). While the genomic complexity is relatively low, mutations in the MEK/ERK pathway and *TNFRSF14* were the most commonly found with a prevalence of 43% and 29% (Louissaint et al., 2016), respectively. Interestingly, *TNFRSF14* mutations were associated with deletion or copy number–neutral loss of heterozygosity of chromosome 1p, which are lesions typically found in *TNFRSF14* of classical FL, evidence of at least a degree of genetic semblance between the two entities. The lack of epigenetic 'addiction' in PTFL along with a low genetic complexity and absence of recurrent mutations found in classical FL has supported PTFL as being considered a biologically and clinically distinct indolent lymphoma.

Duodenal type FL: These are low-grade lesions (grade one or two) localised in the second portion of the duodenum without metastasis (Schmatz et al., 2011). Patients have an excellent prognosis with a 'watch and wait' approach frequently employed. Gene expression reveals similarity to extranodal marginal zone lymphoma of mucosa-associated lymphoid tissue (MALT) lymphoma (Takata et al., 2014). Restricted *V* gene segment usage of immunoglobulins was seen, specifically *VH4* and *VH5*, suggesting an antigen-dependent mechanism in lymphomagenesis (Takata et al., 2009), as in the case of MALT lymphoma, which develops from chronic inflammation. Genetic changes similar to classical FL include the t14;18 translocation and BCL2 and BCL6 expression (Takata et al., 2014). Another similarity to FL is that clones have ongoing somatic hypermutation (SHM) in their *V* genes. However, this is in spite of a lack of AID expression (Takata et al., 2009), suggesting a distinct pathway to facilitating ongoing SHM.

1.6 Transformation

A significant disease event that can occur at relapse or progression is the transformation of FL into an aggressive lymphoma (t-FL) which is considered one of the most unfavourable events in the natural course of FL. t-FL was first described in 1942, in which serial biopsies revealed a 'less differentiated' appearance compared to the preceding FL biopsy (Gall and Mallory, 1942), sharing a more similar histologic appearance to a different lymphoma group. However, the clonal relationship between disease events can be seen through identical IGH-BCL2 rearrangements (Lossos et al., 2002), distinguishing *bona fide* transformation events from *de novo* secondary malignancies. Unlike FL, t-FL has a high proliferative index which can be visualised through Ki-67 staining (Natkunam et al., 2004). The follicular architecture is effaced by the infiltration of diffuse large cells within the follicles (Lossos and Gascoyne, 2011). The majority of transformations resemble de novo diffuse large B cell lymphoma (DLBCL), with a smaller group (7%) resembling Burkitt-like lymphoma (Al-Tourah et al., 2008). The majority of patients with a DLBCL-type transformation have a genetic signature that resembles *de novo* germinal centre B cell type DLBCL (GCB-DLBCL), including BCL2 rearrangements and mutations in EZH2 and TNFRSF14 (Pasqualucci et al., 2014). Intriguingly, 16% of cases are of the activated B –cell subtype (ABC) which are associated with t14;18 negative FLs (Kridel et al., 2015a).

Biopsy and immunophenotyping remain the gold standard for t-FL diagnosis although a combination of clinical features can offer a reliable diagnosis if biopsy sampling is not an option. This includes the sudden onset of B symptoms (fever, night sweats, and weight loss), elevated serum LDH, hypercalcaemia, development of extranodal disease and rapid tumour growth.

The rate of transformation is 2-3% per year (Ban-Hoefen et al., 2013, Wagner-Johnston et al., 2015, Al-Tourah et al., 2008) with a broad incidence rate ranging from 11-70% (Hubbard et al., 1982, Montoto et al., 2007b, Al-Tourah et al., 2008, Garvin et al., 1983, Risdall et al., 1979). This variability is likely due to a lack of uniform across studies, including the cohort size, method of diagnosis by the clinician (histologic versus clinical parameters), variable inclusion of autopsy samples and variable duration of patient

follow up and surveillance. Furthermore, lack of clonality determination using a pre-transformation sample in earlier studies cannot distinguish true transformations from de novo secondary malignancies, impacting the frequency of transformation reported. Identifying patients who are more at risk of transformation remains an attractive line of enquiry for tailored and pre-emptive therapy strategies. Advanced stage disease and a high FLIPI score at diagnosis are predictive of patients at a higher risk of transformation (Montoto et al., 2007b). Furthermore, transformation rates at 5 years following diagnosis were shown to be highest in patients who were observed compared to those who underwent rituximab monotherapy (P = 0.21) (Link et al., 2013), suggesting that initial management strategies during FL may influence transformation risk. However, two different studies contradicted these observations by finding a W&W approach had either a lower or similar transformation risk compared to rituximab therapy (Conconi et al., 2012, Ardeshna et al., 2014). Several studies have looked into whether transformation risk increases over time, with contradicting findings. Montoto *et al* (Montoto et al., 2007b) and Bastion *et al* (Bastion et al., 1997) found that the risk of transformation plateaus after 15 years and 6 years follow up, respectively, suggesting a subset of patients who will never undergo a transformation. However, other studies found that the risk of transformation continues to increase over time, suggesting transformation as an inherent evolution of FL, regardless of treatment versus non-treatment of FL (Horning and Rosenberg, 1984, Al-Tourah et al., 2008). However, both these studies did not include rituximab therapy and therefore the risk of transformation in the rituximab era warrants further examination.

The prognosis of transformed patients was extremely dismal with a median OS of 1-2 years following diagnosis. However, the introduction of rituximab as part of post transformation therapy has seen a considerable improvement with median survival reaching up to 5 years post diagnosis (Alonso-Alvarez et al., 2017, Ban-Hoefen et al., 2013, Link et al., 2013). R-CHOP therapy is used as a treatment option, especially in anthracycline naïve patients, with OS of t-FL patients matching that of *de novo* DLBCL patients on the same regimen (Link et al., 2013, Gleeson et al., 2017). Autologous stem cell transplant (ASCT) in the rituximab era has also shown benefit to patients. For ASCT patients ≤60 years, the 2 year OS was 74% compared to 59% in the non-ASCT patients

≤60 years of which 56% were on R-CHOP therapy (Ban-Hoefen et al., 2013). Despite this improvement, prognostic parameters are still associated with a poorer outcome. For example, patients with a higher FLIPI score, and those who transform less than 18 months following FL diagnosis are indicated to have a poorer survival (Federico et al., 2009, Link et al., 2013). Additionally, patients who are chemotherapy naïve prior to transformation had a superior 2 year OS following ASCT and non-ASCT compared to those exposed to chemotherapy prior to transformation (Ban-Hoefen et al., 2013).

1.7 B cell development

As neoplastic counterparts of normal GC B cells, FL cells are derived from the deviation of the normal B cell development pathway. Therefore, we must understand the processes governing B cell development, as they are implicit to the manifestation of the disease.

B cells are a critical component of the adaptive immune system as they produce antigen-specific immunoglobulins that eliminate pathogens. B cell development begins in the bone marrow where they originate from lymphoid progenitors, which in turn arise from the pluripotent haematopoietic stem cells. The stages of B cell development in the bone marrow are marked by the rearrangement and expression of immunoglobulin genes which occur in a fixed and regulated sequence, as described below.

1.7.1 Immunoglobulin structure and formation

One feature of normal B cells maintained by all FL cells throughout disease is the expression of surface immunoglobulin (sIg). This is in spite of several challenges, including the loss of one functional Ig allele as a result of the t(14:18) translocation and ongoing SHM of the variable regions that increase the probability of stop codon introduction and loss of structural integrity. This selective retention suggests the sIg has an important role in FL pathogenesis.

There are three immunoglobulin gene loci; the heavy chain (IgH) located on chromosome 14q32.3, and two light chains, termed κ and λ, located on chromosome 2p and chromosome 22q, respectively. The protein products of these genes form the mature Ig molecule initially expressed on the pre B cell surface, which in combination with CD79a/b, form the B cell receptor (BCR). The immunoglobulin is made of four polypeptide chains comprising of two identical heavy chains (IgH) and two identical light chains (either Igκ or Igλ) which are held together by disulphide bonds between cysteine residues (Figure 1.1). Heavy and light chains are composed of variable and constant regions. The variable region of the heavy chain is encoded by three gene segments; variable (V_H), diversity (D_H) and joining (J_H). The variable light chain region is encoded by two gene segments; variable ($V_{κ/λ}$) and joining ($J_{κ/λ}$). The variable region of the heavy chains and both regions of the light chains encompass the Fab fragment of the immunoglobulin which determines the antigenic specificity. Antigen binds to 3 complementarity determining regions (CDRs) within this region which are flanked by conserved sequences known as framework regions that act as structural support. The constant region of the heavy chain makes up the Fc portion which is responsible for the effector function of the immunoglobulin dependent on its isotype.

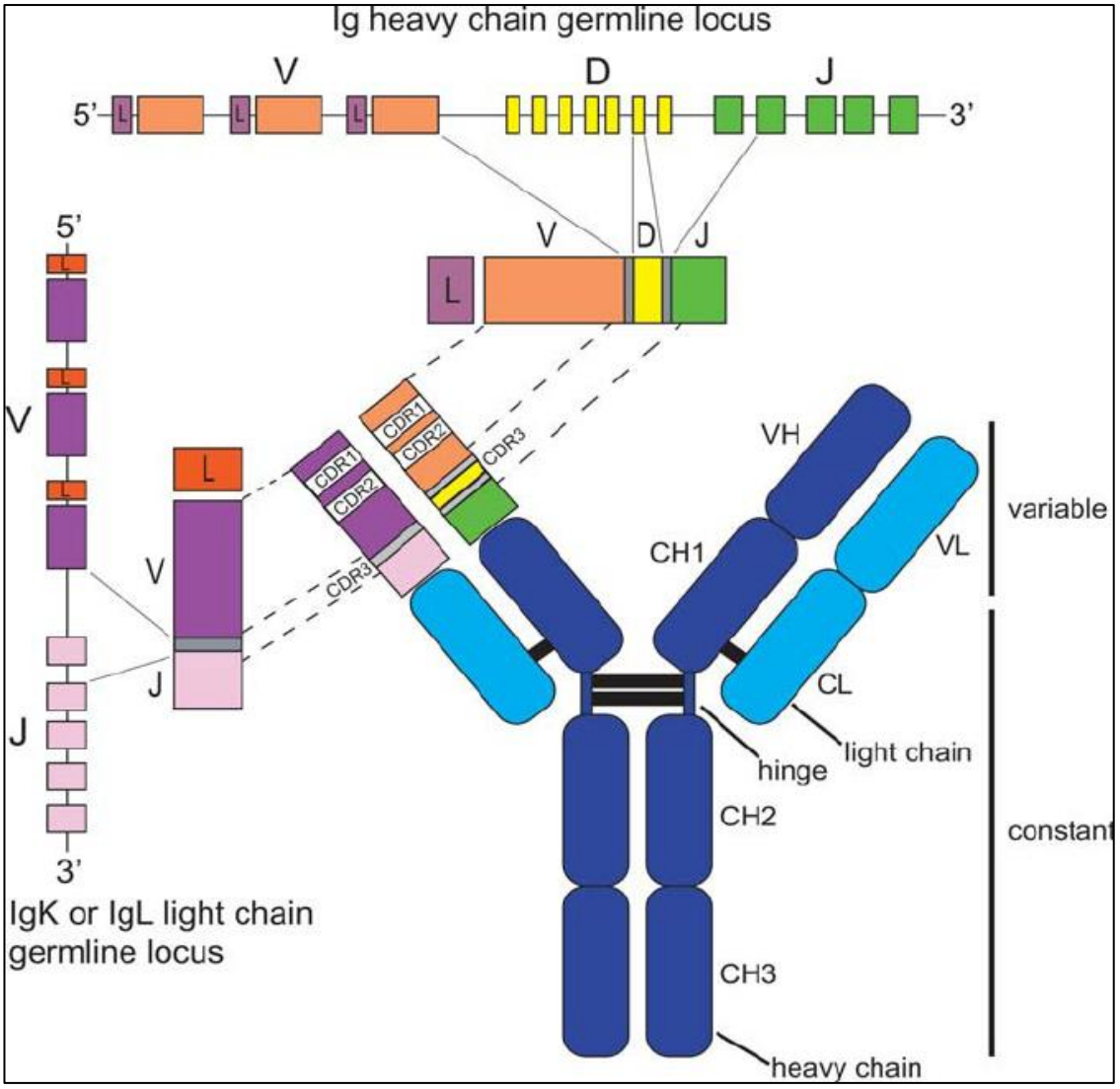


Figure 1.1: *Production and structure of antibody from heavy and light chain germline DNA locus.* Un-rearranged immunoglobulin heavy chain locus (top of figure) contains multiple functional V, D and J segments (of which only a few are depicted) whereas un-rearranged kappa or lambda light chain locus (left of figure) contains multiple V and J segments (of which only a few are depicted). The heavy chain locus rearranges first, with the joining of one V gene segment, one D gene segment and one J gene segment to generate the DNA encoding the heavy chain variable region. This combination is random, contributing to antibody diversity. The light chain V and J segment then combine to generate the light chain variable region. Grey bars represent addition of non-templated bases. The positions of the heavy and light chain variable regions are depicted on the polypeptide, represented by VH and VL, respectively. Complementary determining regions (CDR1-3) are shown, with CDR3 being encoded by all gene segments of the respective light and heavy chains. The constant regions of the heavy and light chains (CH and CL) are encoded by exons (not shown here) that are joined to the rearranged V(D)J gene by mRNA splicing. Disulphide bonds (depicted by black bars) join the heavy and light chains to generate the full antibody structure. Image is taken from Boyd and Joshi (Boyd and Joshi, 2014).

Immunoglobulin production begins during somatic recombination in which gene segments of the variable region are joined together in the bone marrow. The heavy chain (IgH) locus undergoes rearrangement first, with the D and J_H gene segments joining first, followed by the V_H segment joining to DJ_H. This gives rise to μ chains that, in complex with surrogate light chains made of VpreB and delta5, form the pre B receptor that is transiently expressed on the cell surface of large cycling B cell precursors, known as pre-B cells. These cells differentiate into non-dividing small pre B cells in which the light chain genes undergo rearrangement to produce immature B cells that express IgM and low levels of IgD on the cell surface. If the B cell passes IgM autoreactivity testing, it leaves the bone marrow and migrates to secondary lymphoid organs. Chemokines produced by follicular dendritic cells attract the naïve B cells into follicles that are bordered by T cells (T cell zone) and at this stage, B cells are classified as naïve. The B cell receptor binds to a cognate exogenous antigen that enters the follicle and then migrates to the T cell zone where it displays the antigen to a T helper cell as peptides bound to MHC Class II molecules. This enables full activation of the B cell, with some of the B cells differentiating into plasmablasts that produce low affinity antibodies. Other B cells migrate back into the follicle and undergo rapid proliferation to form a germinal centre (GC). Resting B cells are pushed to the periphery of the follicle and form the mantle zone. The germinal centre size continues to increase during proliferation with the establishment of a polarised microenvironment referred to as the dark and light zone. The dark zone refers to the histological appearance of rapidly proliferating B cells being densely packed in contrast to the sparsely populated light zone which also includes a number of other cell types, including follicular dendritic cells, macrophages and T helper cells. The majority of dark zone B cells are termed centroblasts and light zone B cells are termed centrocytes. A majority of B cells undergo apoptosis in the GC with a few being selected to undergo differentiation to either memory or plasma B cells that exit the germinal centre and provide the individual with effective and long lasting protection during chronic or recurring infection. They achieve this by utilising their high affinity BCRs, which are produced in the GC reaction through two mechanisms; somatic hypermutation (SHM) and class switching (CSR) which are discussed in detail below.

1.7.2 Immunoglobulin diversification

To recognise a multitude of antigens and mount an effective immune response requires a diverse antibody repertoire which is generated by four main processes. The first process of immunoglobulin diversity is achieved in the bone marrow. The variable region of each of the immunoglobulin loci contains multiple gene segments. For example, the variable region of the human heavy chain germline DNA contains approximately 40 functional V_H gene segments, 25 functional D_H gene segments and 6 functional J_H gene segments. During formation of the heavy and light chains in the B cell progenitor, only one gene segment from V, D (only in *IGH*) and J are selected for combination. As selection of gene segments is random, a vast number of gene segment combinations can be achieved in different rearrangement events, known as combinatorial diversity. A recombination signal sequence (RSS) adjacent to the coding region of V, D and J gene segments guides recombination. RSSs are composed of three elements; a conserved heptamer contiguous with the coding sequence, followed by a non-conserved spacer region composed of either 12 or 23 base pairs, and a conserved nonamer. The RSS of V and J gene segments contains 23bp spacers and D gene segments have a 12bp spacer. These spacer lengths correspond to approximately one turn (12bp) or two turns (23bp) of the DNA helix. A gene flanked by an RSS containing a 12bp spacer can only join to a gene with an RSS containing a 23bp spacer, known as the 12/23 rule. This ensures that D genes are able to join to V and J genes but V and J genes aren't able to directly join. Recombinases RAG1 and RAG2 bind to the RSS flanking the coding sequences to be joined and through their endonuclease activity, cleaves the RSSs and create a hairpin on the coding end. DNA –dependent protein kinase (DNA-PK) and Artemis catalyse a single-stranded cleavage at a random site near the hairpin, creating a single-stranded tail formed of a few nucleotides from the coding sequence plus a few of their complementary nucleotides derived from the other DNA strand. This generates palindromic or P-nucleotides at the junctions between the V, D and J segments, stretches of nucleotides derived from the complementary strand. Terminal deoxynucleotidyl transferase (TdT) adds non-template encoded nucleotides, or N-nucleotides, to the ends of single-stranded DNA, followed by the pairing of strands, exonuclease trimming of unpaired nucleotides and DNA repair and ligation processes.

This leads to the addition of P- and N-nucleotides at the junctions between gene segments and is a significant contributor to the diversity of the CDR3. As P and N nucleotide addition is a random and unique process to each B cell clone, the composition of the CDR3 region is often employed in clonal analysis studies, including monoclonal B cell malignancies. Once rearrangement is accomplished, the promoter located upstream of the leader sequence in heavy and light chains is brought closer to an enhancer (~ 2kb distance). This close proximity enables the enhancer to activate transcription from the promoter, which is not possible in germline DNA in which the promoters and enhancers are 250-300 kb apart. This ensures that only rearranged immunoglobulins are transcribed to mRNA followed by production of the nascent polypeptide.

A second major source of diversity is achieved through SHM, a process that introduces single point mutations in a step-wise manner throughout the rearranged V region at rates that are 10^6 –fold higher than background mutation rates experienced by other genes. SHM occurs during B cell proliferation in the germinal centre with the overall goal to produce B cells with high-affinity antibodies following antigen activation. SHM occurs during transcription due to the action of the enzyme activation-induced cytidine deaminase (AID) which only acts on single-stranded DNA, in which both strands can be affected. AID is exclusively expressed upon activation of B cells. It deaminates a cytidine to a uridine that has two potential outcomes. Firstly, the uridine can be recognised as a thymidine by DNA polymerases leading to transition mutations (c>t or g>a for the opposite strand). Secondly, the base-excision repair enzyme uracil-DNA glycosylase (UNG) can remove the uridine, generating an abasic site that leads to transition or transversion mutations (c>a or g, g>t or c) through repair by error-prone DNA polymerases. Transition mutations are more frequent than transversions. SHM hotspots exist in the V regions consisting of patterns (a/t)a and g(c/t)(a/t), although SHM can occur anywhere within the region. Mutations can lead to amino acid changes that are silent or replacement. CDRs, which are critical antigen binding sites, tend to accumulate replacement mutations whereas framework regions accumulate silent mutations in order to preserve the overall antibody structure. However, it is important to note that SHM does not discriminate between favourable and non-favourable mutations, leading to several B cell outcomes including 1) antibodies with higher affinity, 2) antibodies with

lower affinity, 3) antibodies with unchanged affinity and 4) antibodies with lost functionality due to premature stop codon introduction. The selection process which occurs in the light zone selects B cells displaying immunoglobulin with higher antigen affinity to undergo more rounds of division and SHM leading to affinity maturation. B cells displaying low affinity or non-functional immunoglobulins undergo apoptosis due to lack of BCR crosslinking and presentation of a peptide to T helper cells. This Darwinian evolution ensures the selection of only B cells displaying high antigen affinity immunoglobulin to undergo differentiation into either plasma or memory B cells.

The third mechanism for antibody diversity is through class switch recombination (CSR). This results in a switched isotype of the heavy chain constant region of an immunoglobulin, leading to a change in effector function without disrupting antigen affinity. Naïve B cells expressing IgM and IgD can change to IgA, IgG, or IgE isotypes, diversifying the antibodies response to a shared pathogen. AID mediates CSR and deaminates a cytosine to uracil within the switch (S) regions upstream of the constant regions (except for IgD). Resulting single-stranded breaks convert into double-stranded breaks within the donor S μ/δ and the acceptor S γ/α/ε. The two S regions are recombined through non-homologous end-joining (NHEJ) with the subsequent deletion of the intervening DNA by loop excision. This results in the replacement of the S μ/δ by the new S γ/α/ε with conservation of the VDJ rearrangement upstream.

1.7.3 B cell receptor signalling

The B cell receptor (BCR) is critical to the development and maturation of B cells and ongoing expression is critical to the survival of mature B cells. Several B cell malignancies, including FL, remain dependent on the expression of the BCR (Kuppers, 2005), suggesting a role in lymphomagenesis.

The surface immunoglobulin forms the extracellular antigen-binding domain of the BCR which is non-covalently bonded to a heterodimer composed of CD79a and CD79b which form the cytoplasmic tail of the BCR. CD79a and CD79b contain conserved peptide motifs called Immunoreceptor Tyrosine-based Activation Motifs (ITAMs) which are

essential for intracellular signalling. Following antigen recognition and engagement with the variable region of the sIg, kinases such as Src family kinases Lyn and spleen tyrosine kinase (Syk) are activated which phosphorylate tyrosine residues in the ITAMs (Yamamoto et al., 1993, Rolli et al., 2002). This, in turn, results in the recruitment of the initial signalling complex composed of Syk, Lyn, Bruton tyrosine kinase (Btk), and adaptor proteins Grb2 and B cell linker (BLNK). Lyn provides continued signal amplification as well as forming a complex with CD19 and other costimulatory molecules that reduce the threshold of B-cell activation and cause BCR aggregation (Fearon and Carroll, 2000). Generation of the signalosome results in the production of second messengers, including diacylglycerol (DAG) and inositol-1,4,5-triphosphate (IP3) (Dal Porto et al., 2004). DAG activates protein kinase C (PKC) whereas IP3 leads to calcium influx from both extracellular compartments and the endoplasmic reticulum (Roos et al., 2005). This, in turn, leads to transcription of NFkB genes which have a broad role in B lymphocyte proliferation, class switching (Ruland and Mak, 2003) and survival (Stadanlick et al., 2008). PKC also activates the MAPK pathway (Hashimoto et al., 1998) which regulates a number of transcription factors which can lead to B cell survival, apoptosis or proliferation. To control BCR activation, negative regulators inhibit activation. These include phosphatases (Pani et al., 1995) and kinases such as Lyn which have both activatory and inhibitory properties (Nishizumi et al., 1998).

BCR signalling underpins the fate of B cells as activation can lead to various functional outcomes including survival, differentiation, proliferation, anergy or apoptosis. This is influenced by signal strength and its modulation by negative and positive co-receptors, in addition to the stage of B cell development and interaction with the microenvironment through secondary signals from non-B cells such as BAFF, CD40, TLRs and IL-4 (Niiro and Clark, 2002). Furthermore, signalling strength can regulate which mature B cell subsets can develop. Peritoneal B cells require the strongest BCR signal, follicular B cells require an intermediate signal and marginal zone B cells require a weaker signal (Loder et al., 1999, Martin and Kearney, 2000, Cariappa et al., 2001). This flexibility and fine-tuning by the BCR ensures the correct cell fate is determined to maintain appropriate populations of different B cell subsets and loss of B cells that are functionally obsolete.

1.8 Immune cells of the germinal centre

The GC is a critical site for B cell antibody affinity maturation during an immune response. This process is dependent on other immune cells in the GC, including follicular dendritic cells (FDCs). FDCs develop from perivascular precursors of stromal cell origin. Their maturation requires stimulation by lymphotoxin and tumour necrosis factor signalling through B cells (Alimzhanov et al., 1997, Pasparakis et al., 1996). FDCs produce CXC-chemokine ligand 13 (CXCL13) that guides B cells that have entered the lymph node towards follicles by signalling through its receptor, CXCR5 (Allen and Cyster, 2008). Therefore, FDCs contribute to maintaining an organised follicular structure. They also promote B cell survival in the GC through production of B cell-activating factor (BAFF) and interleukin 6 (IL-6) (Wu et al., 2009). FDCs display antigen in immune complexes on their surface and are responsible for causing antigen-driven selection of B cells in the light zone, with survival of B cells displaying high affinity antibodies that are able to exit the GC as a memory B cell or plasma cell.

B cell clones which have low affinity antibodies and a lack of sufficient survival signalling undergo apoptosis. Phagocytic cells such as macrophages and dendritic cells clear these apoptotic cells in the GC. Tingible body macrophages are predominantly found in the GC and contain many phagocytized, apoptotic cells in various states of degradation (Smith et al., 1998). While these macrophages are bone marrow derived, it is unclear whether they arise from circulating monocytes or originate during embryonic development as indicated from studies looking into the origin of tissue resident macrophages (Jakubzick et al., 2013, Epelman et al., 2014).

1.9 Early events in FL pathogenesis

FL pathogenesis is a multi-step process, with the likely earliest event occurring in the bone marrow in which an error in VDJ recombination results in the hallmark t(14;18)(q32;q21) translocation. Evidence of pre-malignant cells and lesions suggests a protracted disease course with the accumulation of several genetic hits required for overt disease manifestation.

1.9.1 t14;18 translocation

FL is a germinal centre derived B cell lymphoma with tumour cells avoiding programmed cell death despite being unable to differentiate into plasma or memory B. However, the first considered event in pathogenesis occurs early in B cell development within the bone marrow. 90% of FL patients harbour the t(14:18) translocation in which the *BCL2* locus on chromosome 18q21.3 juxtaposes to the immunoglobulin heavy chain gene (IGH) locus on chromosome 14q32.33 (Cleary and Sklar, 1985). The majority of *BCL2* breakpoints occur within a 150bp region within the 3' untranslated portion of the third exon (~70%), termed the major breakpoint region (mbr). The remainder of breaks occurs 30kb downstream of the *BCL2* locus, termed the minor cluster region (mcr) (Cleary et al., 1986a). *IGH* breakpoints predominantly occur adjacent to the joining elements (JH) or adjacent to the regions where the DH region joins the J region. This suggests that the translocation arises during an error in VDJ recombination, resulting in the *BCL2* gene being under the control of enhancers of the *IGH* locus. This leads to the disruption of one Ig allele and the constitutive expression of the anti-apoptotic protein, bcl2. The t(14:18) mature naïve B cell exits the bone marrow and enters the germinal centre reaction, in which the translocation is likely to exert its pathogenic function. GC B cells normally downregulate bcl2 to enable apoptosis for B cells displaying weak affinity BCRs. Those with high affinity found in the light zone then re-express bcl2 and undergo differentiation. In FL, there is bcl2-driven rescue from apoptosis of B cells with low-affinity BCRs, bypassing a critical mechanism used in normal B cell selection.

1.9.2 t14;18 translocation insufficient for lymphomagenesis

Despite offering apoptotic rescue, the t(14;18) translocation is insufficient to inducing lymphomagenesis. This has been established through the failure of recapitulating the disease in transgenic mouse models overexpressing bcl2. The original model, in which the *BCL2* gene was placed under the control of heavy chain enhancers, developed follicular hyperplasia and after a long period, a high grade lymphoma which was not reminiscent of the indolent nature of FL (McDonnell and Korsmeyer, 1991). The

subsequent mouse model, VavP-*BCL2,* in which the BCL2 gene was placed under the control of the pan-haematopoietic promoter VavP, was able to develop FL disease (Egle et al., 2004). Yet the overexpression of BCL2 in T cells is not found in FL and therefore the molecular features of the disease were not accurately reflected. Around 10% of FL cases are t14;18 negative with the majority not expressing BCL2. Whilst sharing many morphologic and genetic features with t14;18 positive FLs which has maintained their 'classical' FL classification by WHO, there are differences between the two groups. Some negative cases harbour *BCL6* translocations (Horsman et al., 2003) and have micro-RNA expression profiles reminiscent of late GC B cell phenotype (Leich et al., 2011). Furthermore, low frequencies (0.1–10 cells/million) of t(14:18) positive cells are found in peripheral blood mononuclear cells of ~50% of healthy individuals, termed FL-like cells (FLLCs) (Limpens et al., 1995, Dolken et al., 1996, Schuler et al., 2009), without individuals being at an increased risk of FL development. These translocations contain the same breakpoints conserved in FL and results in the overexpression of BCL2, suggesting a similar apoptotic rescue of cells in the GC (Roulland et al., 2006a, Agopian et al., 2009). Interestingly, the number of t14;18 positive cells increase with age and exposure to pesticides (Roulland et al., 2004, Baccarelli et al., 2006, Schuler et al., 2009, Liu et al., 1994), which may support epidemiological findings of a median FL onset at 65 years and increased risk of t14;18 positive NHLs in those with increased use of insecticides (Chiu et al., 2006).

1.9.3 Disease evolution- common progenitor cell compartment

Disease progression had previously been thought of as a linear process, in which the dominant clone of the subsequent disease episode originates directly from the preceding event following the acquisition of additional mutations. This theory was investigated in genetic data gathered from sequential FL biopsies and paired FL/t-FL pairs which offers invaluable insight into FL evolution during progression (e.g. relapse) and transformation. Single nucleotide variants (SNVs) and copy number aberrations (CNAs) from FL-relapse and FL-tFL samples revealed that whilst cases of direct evolution did exist, a greater proportion of patients displayed a non-stepwise evolution pattern,

with the preceding disease event having a set of distinct mutations absent from the subsequent event, suggesting a divergence between the two populations at an earlier time point (Pasqualucci et al., 2014, Eide et al., 2010). A number of shared mutations evidenced a clonal relationship between events. Analysis of the SHM patterns of the variable heavy chain gene in FL-relapse and FL-tFL paired samples revealed relatedness between the two disease events through shared point mutations. However, not all point mutations found in FL were present in the relapse and t-FL samples. This was visualised in genealogic trees composed of IgH-VH sequences, highlighting a complex, non-linear SHM pattern in support of divergent evolution predominance in FL cases (Carlotti et al., 2009, Aarts et al., 2001, Eide et al., 2010). Collectively, these studies imply that successive disease episodes arise from a putative common ancestral cell pool, referred to as the common progenitor cell (CPC) following the acquisition of distinct genetic aberrations (Figure 1.2).

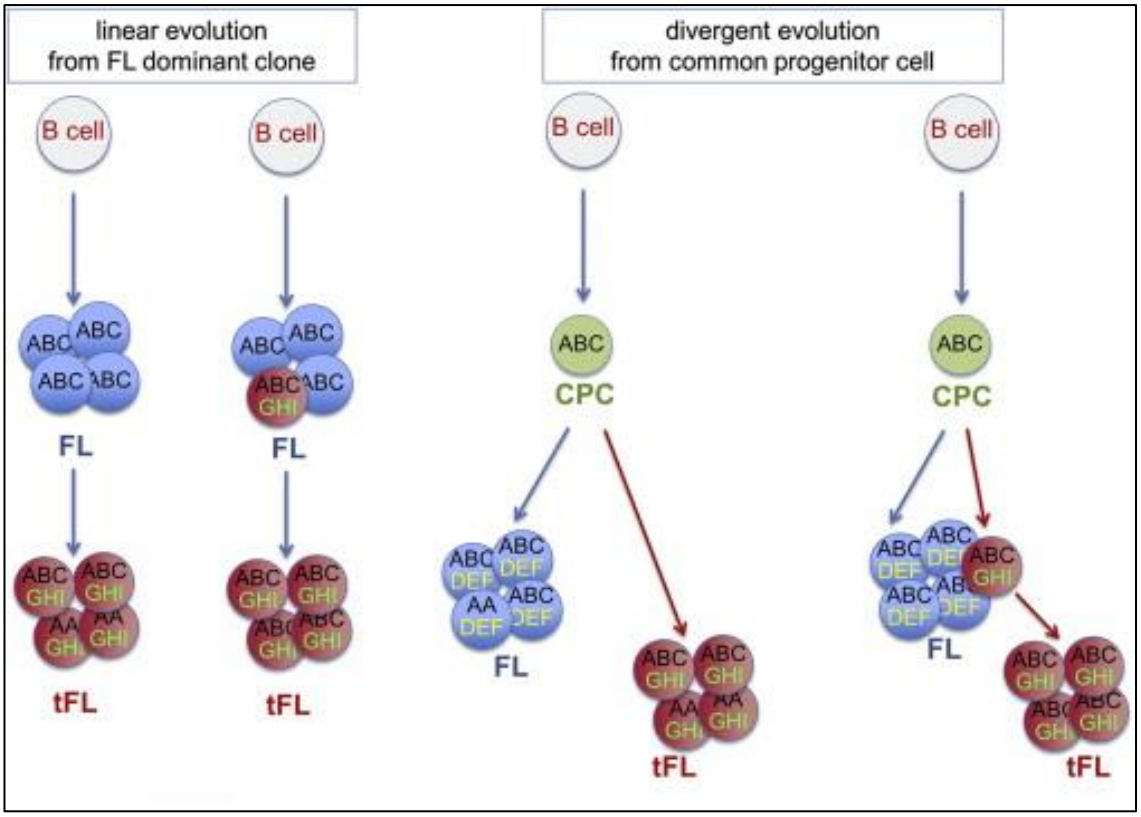


Figure 1.2: *Models of FL evolution as suggested through genomic analysis of sequential FL biopsies and paired FL-tFL samples*. The left-hand panel describes two scenarios of the linear model of evolution, whereas the right-hand panel describes two scenarios of divergent evolution from a CPC. Regarding the linear model, the dominant tFL clone can arise directly from the preceding dominant FL clone following the acquisition of distinct mutations. In the divergent model, both FL and t-FL clones arise from a less mutated, ancestral cell (CPC) following the acquirement of distinct mutations. In both models, the tFL dominant clone may exist as a minor subclone in the preceding FL event which following favourable selection pressures, expand to become dominant. Image is taken from Pasqualucci et al (Pasqualucci et al., 2014).

Studies on two cases of donor derived FL following stem cell transplant has given us insight into the features of this population. In the first case, a son received an allogeneic stem cell transplant from his father for the treatment of acute myeloid leukaemia (AML) (2009). Both father and son developed clonally related FL with transformation 3 and 10 years after transplantation, respectively. Evidence of a CPC was indicated through the *IGHV* sequence of the major clones from the two tumours, with shared and distinct point mutations between the father and son (Carlotti et al., 2009). The long latency between a transfer of the CPC in the stem cell transplant and disease manifestation (up to ten

years) reveals the CPC population has a self-renewal capacity and can remain dormant for years. This long dormancy was further supported by a second case of donor derived FL. Siblings who had underwent allogeneic SCT and donor leukocyte infusions (DLI) for the treatment of chronic myeloid leukaemia (CML) developed FL seven years after treatment (Weigert et al., 2012). The same t14;18 breakpoint and VDJ rearrangement were found in the FL tumours of the siblings, evidencing their clonal relationship. Interestingly one in two thousand cells in the DLI also harboured these features. Additionally and rather remarkably, ultra-deep sequencing of the DLIs identified fourteen of the fifteen shared somatic mutations identified in the tumours. As DLIs were taken at the time of transplant, we can infer that the CPC population existed at least seven years before overt disease manifestation. It also suggests that the CPC is able to circulate in the peripheral blood, taking up residence in different anatomical sites. The pattern of evolution from this putative CPC was investigated through whole genome and exome sequencing of paired FL-t-FL samples and the analysis of non-synonymous mutations across events (Okosun et al., 2014). All cases had shared somatic mutations between disease episodes, indicating the existence of a CPC. Yet the degree of clonal semblance varied between cases, revealing two distinct evolutionary patterns. A high number of shared mutations between disease episodes indicated evolution through a 'rich' CPC whereas a low number of shared mutations indicated a 'sparse' CPC in which disease events diverge at an earlier time point (Figure 1.3). The majority of cases were classified as having a 'rich' CPC, with mutations occurring in histone-modifying genes (*CREBBP*, *KMT2D*) and immune modulating genes (*B2M*, *CD58*, *TNFRSF14*), suggesting that these aberrations are early events in pathogenesis. Over 70% of cases had concurrent mutations in at least two chromatin modifying genes. Several other studies have implied chromatin-modifying genes as an early event in FL (Pasqualucci et al., 2011a, Pasqualucci et al., 2014, Li et al., 2014, Green et al., 2015). The functional consequence of *CREBBP* mutations was shown to result in the decrease of MHC class II expression on tumour cells, leading to immune evasion (Green et al., 2015). Intriguingly, Okosun *et al* observed that different mutations occurred within the *KMT2D*, *TNFRSF14* and *CREBBP* genes of 'sparse' serial biopsies (Okosun et al., 2014). This indicates a convergent evolution in which similar genetic alterations occur independently within episode-specific clones. The observation of alterations in these genes across both

evolutionary patterns suggests they have an important role in lymphomagenesis and progression.

Due to the high relapse rates seen across FL patients, the CPC population is believed to be resistant to current therapies. Targeting early events may lead to the disruption of this reservoir pool improving remission rates and progression free survival. Epigenetic-based therapies are conceptually attractive and would move towards a precision medicine approach to FL treatment that is currently based on a 'one size fits all' approach. However, several trials using epigenetic modifying agents, such as histone deacetylase (HDAC) inhibitors and EZH2 inhibitors as single agent therapies, have shown disappointing complete response rates (Evens et al., 2016, Morschhauser et al., 2017, Kirschbaum et al., 2011, Ogura et al., 2014). Future research should focus on the functional implications of these therapies and exploration of how to improve efficacy, for example through combination regimens.

While sequential biopsies and donor derived FL cases has provided a 'proof of principle' of the CPC compartment, little is known regarding its phenotypic and biologic features as well as its location (in either tissue or peripheral circulation) and stage of differentiation (e.g. ranging from cancer stem cell to an already malignant subclone). Characterising this compartment represents a significant scientific challenge yet one that would provide invaluable insight into disease evolution and opportunities for therapeutic intervention.

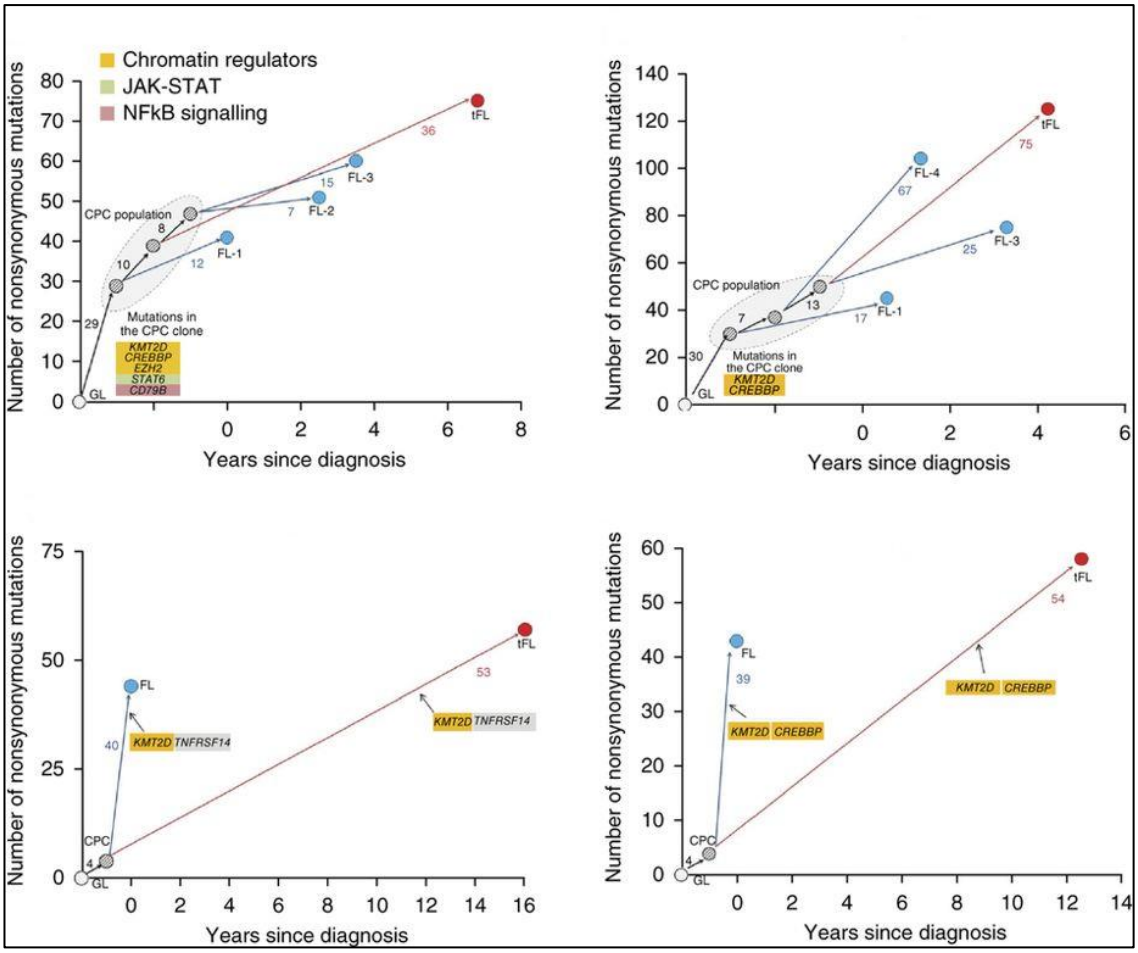


Figure 1.3: *Phylogenetic trees based on non-synonymous mutations of paired FL-tFL biopsies.* Upper and lower panels highlight distinct evolutionary patterns The branches represent individual disease events (blue circles indicative of FL and relapsed FL, red circles indicative of transformation), with the number of event specific mutations indicated on each branch. Trees on the top panel reveal two cases of evolution from a 'rich' CPC population (indicated by grey circles) in which serial samples share a high number of mutations. Trees on the bottom panel reveal two cases of evolution from a 'sparse' CPC, with events having few shared mutations (only 4 in both examples). Image is taken from Okosun *et al* (Okosun et al., 2014).

1.9.4 Intermediate steps in FL development

The long latency period to disease presentation seen in allo-HSCT derived FL cases, suggests FL development is protracted, taking years to develop the disease. This ties in with the multi-step pathway to lymphomagenesis, requiring several genetic hits in addition to the t14;18 translocation. Whilst studies have suggested epigenetic alterations to be early events in disease evolution, the exact steps and time points at

which they are acquired is still unknown. However, now we have begun to identify potential precursor cells and lesions in which such genetic mutations may arise and accumulate in to develop overt disease. The current model of FL pathogenesis from these presumed precursor lesions is summarised in Figure 1.4.

Although circulating t14;18 positive cells termed follicular lymphoma-like B cells (FLLCs) are found in healthy individuals who never go on to develop the disease, they harbour features which are reminiscent of FL cells including expression of the GC B cell marker CD10, and upregulation of BCL2. Furthermore, FLLCs have intraclonal diversity (Agopian et al., 2009), suggestive of repetitive GC entry where they can undergo AID-mediated SHM and/or CSR. Indeed, FLLCs have been shown to have a GC-experienced memory B cell phenotype (Roulland et al., 2006b, Hirt et al., 2007), supporting their increased exposure to AID. One study also found that some FLLCs have aberrant ongoing AID activity which is normally downregulated following GC exit (Agopian et al., 2009). As AID introduces aberrant mutations in non-Ig genes which are usually proto-oncogenes, such as *BCL6*, *PIM1*, MYC and *PAX5* (Rossi et al., 2005, Pasqualucci et al., 2001, Liso et al., 2006) increased exposure may increase genomic instability, leading to accumulation of mutations in FLLCs and their eventual transformation to overt FL.

The direct relationship between FLLCs and FL was established in a study which found that the frequency of FLLCs can be used as a predictive marker for FL development. Pre-diagnostic blood samples were available for 100 FL patients (Roulland et al., 2014). Samples were available due to the patients being involved in an earlier study before the occurrence of overt disease. t14;18 positive cells were screened using sensitive PCR assays and clonal relationships between FLLCs and tumour samples were assessed by analysing the t14;18 breakpoints. All FLs and identified t14;18 positive cells in the paired pre-diagnostic blood sample were clonally related, establishing the FLLCs as the disease precursor cell. Furthermore, pre-diagnostic blood samples from these patients had a higher frequency of t14;18 positive cells when compared to samples taken from a control group who didn't go on to develop the disease (n=218). They concluded that individuals with t(14;18) frequency reaching 1 in every 10,000 blood cells had a 23-fold greater risk of progression to FL. This predictive value did not change depending on

whether FLLCs were harvested close to disease onset or many years before. This highlights the long latency period between pre-malignancy to diagnosis, suggesting the overt disease is preceded by a preclinical, asymptomatic phase of clonal growth and establishment of a precursor clone that takes years, if not decades to manifest. This supports the hypothesis of a long-lived and dormant CPC population. This was also seen in cases of donor-derived FL, in which t14;18 positive cells detected at transplantation within DLIs contained the same *BCL2:IGH* rearrangement as that found in the FL clone of the donor and recipient years later (Weigert et al., 2012).

t14;18 positive cells are also found incidentally in GCs of 2-3% of reactive lymph nodes of healthy adults, forming unconventional foci within a background of BCL2 negative B cell follicles (Cong et al., 2002, Henopp et al., 2011). They are monoclonally derived and colonise several GCs. They are termed *in situ* follicular neoplasia (ISFN), formerly known as FL *in situ*. In a study of 21 patients with FLIS, <5% progressed to a clonally related FL (Jegalian et al., 2011). However, the study was limited by the small cohort size and a restricted median follow up 118 months. A study by Mamessier et al utilised high-resolution comparative genomic hybridization microarrays on laser-capture micro-dissected ISFN cases to reveal that there is a degree of genetic instability in these lesions, albeit at a much lower number when compared to overt FL (Mamessier et al., 2014). A study by Schmidt utilised paired FL-ISFN samples to analyse the genetic relationship between FL and its pre-malignant lesion (Schmidt et al., 2014). ISFN samples showed little to no secondary genetic changes. Only one case had an EZH2 mutation shared between biopsies. The low genetic complexity and semblance of ISFN with clonally related FL highlights that several intermediate steps are necessary for disease manifestation from precursor cells.

As FLLCs and ISFN are considered precursors in lymphomagenesis and express t14;18, we can infer that the survival advantage offered by constitutive expression of BCL2 enables the B cell to acquire and accumulate genetic aberrations over the years, if not decades, to enable malignant transformation. The relationship between FLLCs and ISFN remains undetermined, however, Cheung *et al* found that in one ISFN case, a paired blood sample contained B cells with an identical t14;18 rearrangement to the ISFN cells

(Cheung et al., 2009b). This supports the hypothesis that FLLCs represent the circulating counterpart of ISFNs. The selection and commitment of some ISFN to develop into FL whilst others do not suggest a diverse number of unknown factors at play in contributing to lymphomagenesis.

Whilst in ISFN the lymph node architecture is preserved, in partial involvement by FL (PFL), the affected follicles are larger with ill-defined marginal zones and attenuated mantle cuffs. Therefore, histologically, PFL has a greater semblance to FL. This is supported in a follow-up study that found that 53% of untreated PFL patients went on to develop overt FL over a fourteen year period (Jegalian et al., 2011) compared to only 5% of FLIS patients. PFL also has significantly more genetic alterations compared to FLIS as evidenced in array comparative genomic hybridization studies (Mamessier et al., 2014). Some of the amplified genes were also found to be amplified in low-grade FL, indicating a selective pressure for the high malignant transformation of PFL patients. These recurrent mutations included *TNFRSF14*, *EZH2*, and *MLL2*. However it's important to note that genomic alterations and gains/losses within the PFL cases was still low compared to overt FL and that most cases present with low stage disease, suggesting that PFL does not simply represent a partial colonisation of the lymph node by FL, but is rather a precursor lesion (Adam et al., 2005).

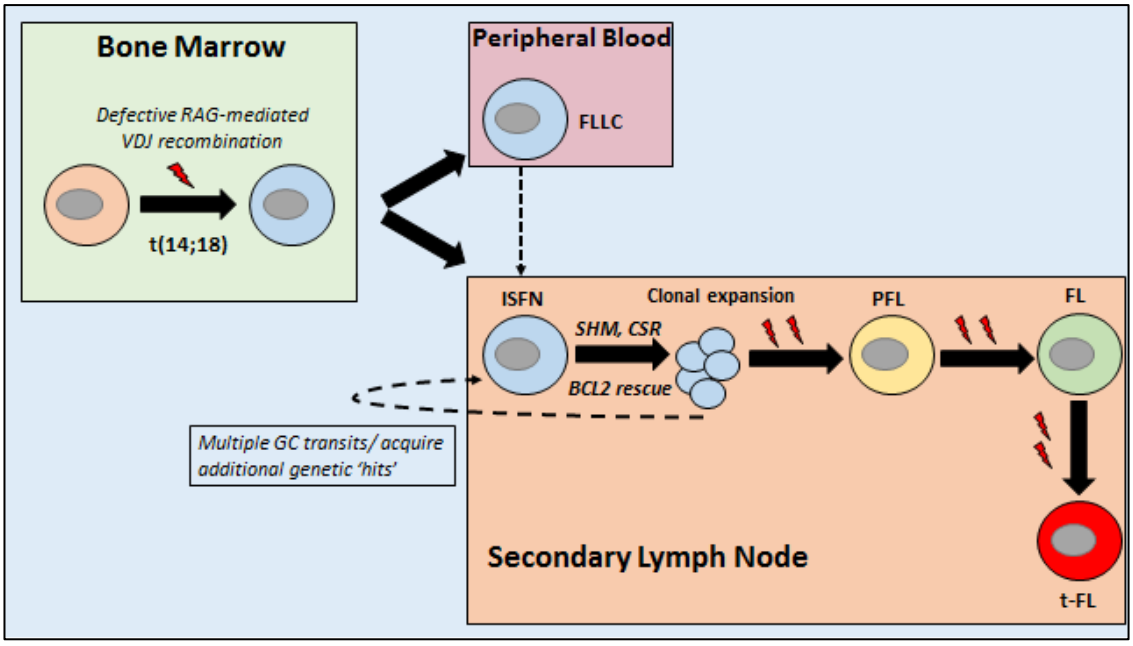


Figure 1.4: *Model of FL pathogenesis based on presumed t14;18 positive precursor cells*. In the bone marrow (BM), naïve B cells acquire the t(14:18) translocation as a result of defective VDJ recombination. Cells exit the BM and enter peripheral circulation in which cells are termed FLLC. Cells colonise secondary lymphoid tissues where they undergo the GC reaction, characterised by somatic hypermutation (SHM) and class switch recombination (CSR). Constitutive expression of BCL2 protein provides a survival advantage, rescuing cells from apoptosis regardless of BCR affinity. ISFN is believed to represent this pre malignant precursor stage. Cells undergo clonal expansion within the GC and can exit into peripheral circulation. They are able to re-enter the GC multiple times as seen by FLLCs presenting a GC-experienced memory B cell phenotype (Roulland et al., 2006b, Hirt et al., 2007), where they have opportunity to increase their genomic instability through repeated exposure to AID. As PFL has a greater genomic complexity than ISFN (Mamessier et al., 2014), they are believed to represent a more differentiated precursor cell. The gain of specific genetic and epigenetic hits (↯) then leads to overt disease and t-FL.

1.10 Immunoglobulin in FL

One feature of normal B cells maintained by all FL cells throughout disease is the expression of surface immunoglobulin (sIg). This is in spite of several major challenges, including the loss of one functional Ig allele as a result of the t(14:18) translocation and ongoing SHM of the variable regions that increase the probability of stop codon introduction and loss of structural integrity (Zuckerman et al., 2010). This selective retention suggests the sIg has an important role in FL pathogenesis. Further evidence of a conferred selective advantage came from the observation that resistance to anti-

idiotypic antibody therapy directed against the Ig variable region resulted in the outgrowth of tumour cells with amino acid substitutions in the targeted region rather than the outgrowth of Ig negative variants (Cleary et al., 1986b).

As a monoclonal disorder, FL cells can be identified through identical *V*, *D* and *J* genes within their variable region, established during somatic recombination in the bone marrow. The usage of *IGHV* genes resembles the repertoire of normal peripheral B cells with *IGHV*3-23 the most commonly used gene (Berget et al., 2015, Brezinschek et al., 1997, Aarts et al., 2001). *IGHD3–10* and *IGHJ4* are the most commonly used *D* and *J* genes used, respectively. As *IGHV3* is the largest subgroup family, it is unsurprising that the majority of FL patients belong to this group, suggesting an unbiased usage of *IGHV* in FL. This is in contrast to chronic lymphocytic leukaemia (CLL) which unlike normal B cells, has a notable over-representation of genes, including *IGHV1-69* and *IGHV4-34* (Schroeder and Dighiero, 1994, Fais et al., 1998) indicative of a superantigen contributing to disease pathogenesis. The usage of *IGHV* genes has also been shown to be a prognostic marker in disease. For example in CLL, *V3-21* is associated with an unfavourable prognosis (Ghia et al., 2008, Oscier et al., 2010). A novel finding by Berget et al found that FL patients with *IGHV5* usage had an unfavourable prognosis with an estimated survival at five years of 62.5% compared to patients in all other subgroups who had a combined five year survival of 95.1% (Berget et al., 2015). However, validation in a larger series is warranted for validity of *IGHV* sequence analysis in a prognostic setting.

In accordance to their derivation from GC B cells, FL cells accumulate mutations in their variable region genes through SHM. Ongoing somatic hypermutation, compatible with their prolonged expansion in the GC (Kuppers et al., 1999), results in intraclonal sequence heterogeneity of malignant clones (Aarts et al., 2000) which can be used to trace the clonal evolution of the tumour (Bognar et al., 2005). In normal GC cells, SHM results in modifying the affinity of the BCR for its cognate antigen, resulting in either the proliferation/expansion of the B cell clone through strong antigen engagement, or its clonal deletion due to loss of survival signalling (Klein and Dalla-Favera, 2008, De Silva and Klein, 2015). Early reports of the tumour *IGHV* genes having a similar SHM pattern to those of normal antigen-selected B cells (77, 78), supported an antigen-driven

selection and expansion in malignancy. This included replacement mutations in the CDRs of the BCR, and an underrepresentation of replacement mutations in the framework regions (Jacobs and Bross, 2001, Bognar et al., 2005). However, findings that are more recent are proving contradictory, with no positive selection in the CDRs of FL Ig variable regions but a strong negative selection against replacement mutations in the FWRs (Hershberg et al., 2008, Zuckerman et al., 2010). This suggests that based on SHM patterns, FL clones are selected for their ability to express a functional BCR and not for antigen recognition.

The majority of FL cases express sIgM, with ~40% having undergone isotype switch to IgG (Scherer et al., 2016, Vaandrager et al., 1998). Intriguingly, CSR frequently occurs in the translocated non-functional IGH allele of IgM cases (Vaandrager et al., 1998, Akasaka et al., 1998). In normal memory B cells, CSR only occurs on the functional allele, making this observation a hallmark feature of FL cells. In FL, CSR occurs as frequently on both alleles, but "downstream switch" (e.g., Sγ-to-Sα) happens at unusually high frequency on the productive allele, sparing the Cμ region from deletion and allowing sIgM expression (Figure 1.5) (Vaandrager et al., 1998). Consequently, despite CSR occurring on both alleles in >80% of cases, most FL cases still express a sIgM. This 'allelic paradox' indicates the presence of a selective pressure in favour of sIgM expression on a B cell population that is at the same time permanently driven to switch. This could be due to favourable downstream effector functions offered by the cytoplasmic tail of IgM. Interestingly, t14;18 positive cells in healthy individuals also display this 'allelic paradox' (Roulland et al., 2006b), highlighting a possible shared functional feature between malignant cells and their precursors.

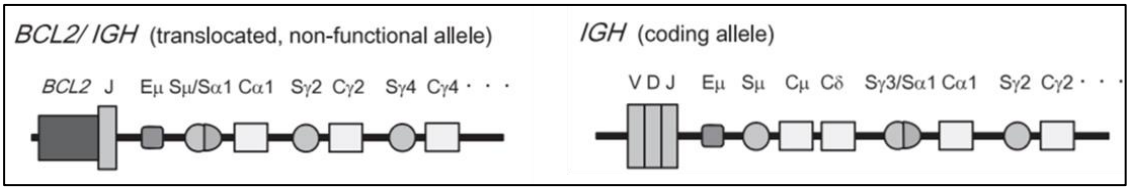


Figure 1.5: *CSR processes in translocated and functional IGH allele.* A high frequency of CSR are observed in the translocated allele. In contrast, the C μ region is usually spared from deletion in the coding IGH allele, while "downstream switching" is observed at high frequency (left panel). Image adapted from Kishimoto and Nishikori, 2014 (Kishimoto and Nishikori, 2014)

1.11 B cell receptor signalling in FL

Due to evidence of retention of the Ig, as described above, it is likely that BCR signalling is important to tumour survival. Krysiak *et al* found that ~40% of FLs had recurrent mutations affecting the BCR signalling pathway including CD79B, CARD11, CXCR4, and BTK mutations in the tyrosine kinase domain (Krysiak et al., 2017). Lower level mutations were also found in *BLNK*, *PLCγ2*, *BCL10*, and *NFκB2*.

The investigation into BCR kinetics and activation in normal B cells and FL B cells can help provide insight into potentially important signalling differences which may contribute to pathogenesis. An interesting study by Irish *et al* investigated phosphorylation of BCR effector proteins in malignant B cells and non-malignant B cells derived from both healthy individuals and those within the FL tumours. Basal phosphorylation of Btk, Syk Erk1/2 and p38 did not differ between malignant and non-malignant B cells, disproving the hypothesis of constitutive activation of BCR signalling in FL (Irish et al., 2006). Another unexpected finding was that following BCR cross-linking, phosphorylation was lower in FL cells although Erk1/2 phosphorylation was variable between patients with some having increased phosphorylation and others having reduced phosphorylation. As this was independent of Syk and Btk phosphorylation, it suggests a novel signalling route activated between BCR and Erk1/2 in FL. When BCR cross-linking was repeated in the presence of hydrogen peroxide (H_2O_2), a phosphatase inhibitor, there was increased phosphorylation amongst all four phosphoproteins, suggesting BCR signalling in FL has a strong negative regulation by protein tyrosine phosphatases (PTPs). Compared to infiltrating non-tumour B cells, BCR-mediated signalling via phosphorylation of Btk, Syk, Erk1/2, and p38 occurred more rapidly in tumour B cells and sustained greater levels of per-cell signalling which was sustained for hours. There was strong inter-individual variability in both the magnitude and kinetics of the signal, highlighting the heterogeneity in BCR signalling in FL disease. Heterogeneity was further supported by a later study by Irish *et al* which identified an FL cell subset with impaired BCR signalling that was associated with an unfavourable prognosis (Irish et al., 2010). Phosphatase inactivation reversed BCR signalling inhibition in these cells, suggesting increased negative regulation as a cause of impaired signalling rather than an abrogation of the

BCR signalling pathway. As the disease progressed following therapy, the subset expanded, highlighting the changes to BCR signalling as the disease evolves and different subclone populations gain dominance. This expansion could support the notion of intra-tumour signalling heterogeneity contributing to therapy failure and relapse.

Amin *et al* also compared BCR signalling in FL cells, looking specifically at IgM+ and IgG+ FL cases. They compared findings to healthy IgM+ or IgG+ GC B cells. IgM+ FL cells were shown to be effectively stimulated through cross-linking alone when compared to their normal GC B cell counterpart, despite having significant phosphatase activity (Amin et al., 2015). However, IgG+ FL cells did not exhibit an increased BCR response following cross-linking alone compared with normal IgG+ GC B cells and was even shown to have reduced signalling which was corrected following phosphatase addition. This heterogeneity in signalling could be due to reduced BCR expression on IgG+ FL cells and different phosphatase activity between the two isotype groups.

Regarding the source of BCR stimulation, two scenarios have been hypothesised. The first is through antigen stimulation. A study in which tumour derived Ig was tested against tissue antigens revealed that 26% of FLs were self-reactive (Sachen et al., 2012). Myloferin was identified as a self-antigen for one patient and despite ongoing SHM, self-reactivity of the BCR remained constant in tumour subclones, suggesting a positive selection of a BCR that recognises the antigen. Building on this, Cha *et al* found that a subgroup of patients (19.35%) had immunoglobulins that were reactive to vimentin, indicating it as a shared autoantigen (Cha et al., 2013). Interestingly, these cases were more commonly IgG+ than IgM+, suggesting that IgG+ FLs are more commonly self-reactive than IgM+ cases. This observation, in addition to only a subset of FL cases being self-reactive, suggests there is heterogeneity regarding the source of BCR stimulation.

The second inferred source of BCR activation stems from the observation that the majority of FLs carry high mannose glycans in their Ig variable region (Radcliffe et al., 2007). These bulky oligosaccharides impair BCR affinity to its cognate antigen (Schneider et al., 2015), however, activate BCR signalling through binding to calcium-dependent (C-type) lectins which are found in the tumour microenvironment (Coelho et al., 2010, Linley et al., 2015). This antigen-independent BCR activation pathway will be discussed

in depth in the chapters below. Figure 1.6 summarises the two proposed sources for BCR stimulation.

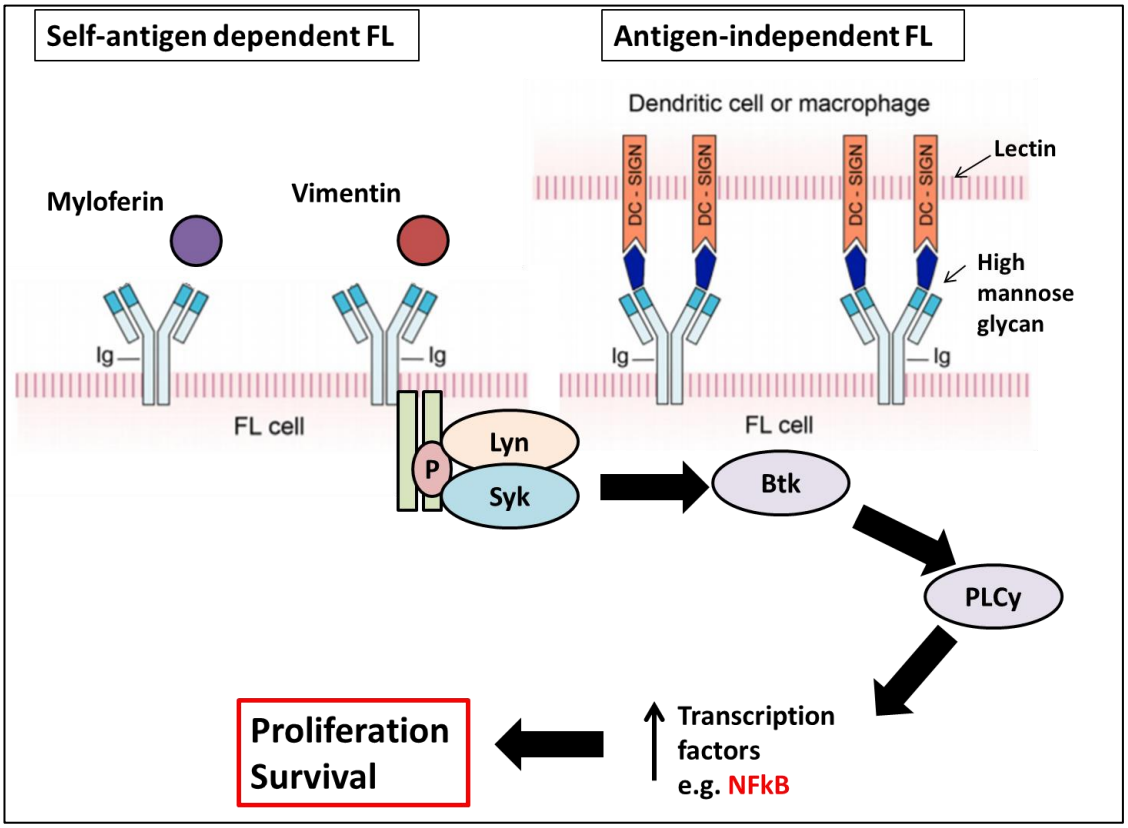


Figure 1.6: *Model of BCR stimulation in FL*. The left-hand panel highlights FL cases which are capable of recognising self-antigen. Two identified antigens, myloferin and vimentin, are shown. The right-hand panel depicts the scenario of high mannose glycans in the variable region of the immunoglobulin that are able to interact with lectins expressed on immune cells, such as dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin (DC-SIGN) and activate BCR signalling through an antigen-independent pathway.

1.12 Genetics of FL

Next-generation sequencing (NGS) platforms have begun to unravel the genetic landscape of FL tumours by identifying several recurrent somatic mutations in genes involved in a variety of cellular pathways, including immune surveillance and signalling pathways associated with BCR, NFKB and JAK-STAT. Table 5 gives a non-exhaustive list of genes recurrently mutated in patients. Mutations can arise through loss or gain of

chromosomal segments, with certain alterations being more frequently found in FL, as described below.

Aberrations in histone modifying genes are present in nearly all FL patients, indicating epigenetic deregulation as a hallmark of FL. Epigenetics, which means 'outside conventional genetics' is used to describe modifications of gene expression and activity without affecting the DNA sequence (Felsenfeld, 2014). This includes DNA methylation and histone modification.

Pathway	Gene	Frequency (%)	Function	Oncogenic alteration
Epigenetic regulation	*KMT2D*	70-90	Histone H3K4 methyltransferase	Loss of function
	CREBBP	50-79	Histone H3K27 and H3K18 acetyltransferase	Loss of function
	EZH2	20-30	Histone H3K27 methyltransferase	Gain of function
	EP300	10-20	Histone H3K27 and H3K18 acetyltransferase	Loss of function
Transcription factors	*STAT6*	10-15	JAK-STAT signalling	Gain of function
	FOXO1	10	Downstream of BCR signalling	Gain of function
	MEF2B	10-20	Regulator of *BCL6*	Gain of function
Signalling	*RRAGC*	15-20	Guanine nucleotide-binding protein	Gain of function
	TNFRSF14	30-40	Inhibit HVEM signalling	Loss of function
	CARD11	10-15	BCR–NF-κB signalling pathway	Gain of function
Nucleosome remodelling	*ARID1A*	10-15	Alter chromatin structure	Unknown
	BCL7A	10	Alter chromatin structure	Unknown

Table 1.5: *Recurrent mutations in FL with frequencies of 10% and above*. Information was adapted from Kuppers and Stevenson (Kuppers and Stevenson, 2018) and Huet *et al* (Huet et al., 2018a).

1.12.1 Chromosomal alterations: 1p36 and 6q deletions

In addition to the t14;18 translocation, FL is characterised by additional chromosomal aberrations, with the most common being non-random losses of 1p36 and 6q and gains of 7, 18, and X (Schwaenen et al., 2009, Johnson et al., 2008, Cheung et al., 2009a, Horsman et al., 2001). 1p36 region alterations (deletions and acquired uniparental disomy) were described as one of the most frequent secondary genetic abnormality in FL (Cheung et al., 2010a), with 6q deletions found in 10-15% of cases (Cheung et al., 2009a). NGS has begun to identify genes affected by these aberrations. The tumour necrosis factor receptor superfamily 14 (*TNFRSF14*) gene is located on 1p36, and following a ~12kb deletion of the region, is inactivated (Cheung et al., 2010a, Launay et al., 2012). Mutations in *TNFRSF14* occur at high frequencies in FL patients (30-40%) and in PTFL patients (>25%) (Martin-Guerrero et al., 2013, Louissaint et al., 2016, Schmidt et al., 2016) which otherwise have very contrasting genetic landscapes. TNFRSF14 lesions are similar between the two variants, including deletion or copy number–neutral loss of heterozygosity of chromosome 1p. *TNFRSF14* encodes for the herpes virus entry mediator (HVEM), a transmembrane receptor expressed on B cells that attenuates T cell proliferation and effector response following engagement with T cell expressed B- and T-lymphocyte attenuator (BTLA) (Sedy et al., 2005, Gavrieli et al., 2003, Krieg et al., 2005, Watanabe et al., 2003). The HVEM-BTLA axis was also shown to inhibit BCR signalling by reducing Syk activity and in turn reduced activation of downstream signalling molecules (Vendel et al., 2009). Knockdown of HVEM or BTLA in the VavP-BCL2 mouse model resulted in accelerated progression of FL development and additionally resulted in a tumour supportive microenvironment, through an increase in TNF family cytokines and follicular T helper cells (T_{FH}) which secrete IL-4 and IL-21 (Boice et al., 2016, Crotty, 2014). This highlights HVEM as having a tumour suppressor function, which is supported by its ability to inhibit proliferation of adenocarcinoma cells (Harrop et al., 1998) and enhance Fas-induced apoptosis in non-Hogkin's lymphoma cells (Costello et al., 2003). However, how *TNFRSF14* mutations translate to clinical outcome is unclear. A study by Cheung *et al* found that *TNFRSF14* mutations correlated with an inferior OS and disease-specific survival (DSS) (Cheung et al., 2010b). In contrast, Launay *et al* found mutated patients had a better overall survival than non-mutated (Launay et al., 2012). However, HVEM

can mediate several costimulatory and co-inhibitory pathways (Shui et al., 2011) and therefore further studies are necessary to determine the global functional consequences of TNFRSF14 mutations in FL.

The second most common chromosome alteration is the deletion of 6q. Several studies have shown that 6q deletions are linked to patient survival suggesting the presence of tumour suppressor genes within this region (Offit et al., 1993, Nanjangud et al., 2007). A seminal study by Oricchio *et al* used a pooled shRNA library specific for genes located in the different 6q common regions of deletion and then tested their ability to rescue BCL2 overexpressing pro B-lymphocytes following growth factor withdrawal (Oricchio et al., 2011). This study provided the first evidence for Ephrin receptor A7 (*EPHA7*) as a tumour suppressor gene. The knockdown of EPHA7 in the FL mouse model, vavPBcl2, resulted in accelerated lymphomagenesis, similar to the effects of *MYC*. FL histology was unchanged and lymphomas had a low Ki-67, resembling the slow proliferation seen in human FL. When recombinant EPHA7 ectodomain (EPHA7FC) was administered to lymphoma cell lines, the proliferation of cells was dramatically reduced, highlighting EPHA7's role as a tumour suppressor in FL pathogenesis. This was in part the result of EPHA7 inhibiting another Ephrin receptor, EPHA2, from forming a homodimer and its subsequent oncogenic signalling through ERK and STAT3. When FL TMAs were examined for EPHA7 expression, 72% were completely absent for the protein in contrast to the clear expression seen in healthy tonsil controls.

The impact of chromosomal alterations on clinical outcome, by themselves or in combination, has been investigated (Cheung et al., 2009a, Horsman et al., 2001, Viardot et al., 2003). While these studies have suggested a correlation between aberrations and inferior survival, findings were highly variable likely due to a lack of experimental reproducibility and patient variability. A later study by Johnson *et al* (Johnson et al., 2008) performed on diagnostic samples exclusively found that an additional X chromosome in males was the only aberration to correlate with prognosis, indicating an inferior overall survival.

1.12.2 Epigenome

While epigenetic modifications are essential to mammalian development (Reik, 2007), its deregulation can result in altered transcription states, contributing to a malignant transformation in both solid and hematologic tumours. This is expressed through global DNA hypomethylation and silencing of tumour suppressor genes through hypermethylation, resulting in genomic instability (Esteller, 2008, Feinberg and Vogelstein, 1983, Merlo et al., 1995). Epigenetic deregulation is considered a hallmark of FL pathogenesis, given that mutations in chromatin-modifying genes occur with high frequency and often occur concurrently with other epigenetic modifiers within individual cases, suggesting that promotion of lymphomagenesis through epigenetic deregulation is the result of a culminated effect. Mutations affect the function of several histone methyltransferases and histone acetyltransferases, including *KMT2D* (also known as *MLL2*) and *CREBBP*, respectively. Truncating and missense mutations in *KMT2D* are found in 60-90% of FL patients whereas loss of function mutations in *CREBBP* are found in up to 70% of patients (Green et al., 2013, Morin et al., 2011, Okosun et al., 2014, Pasqualucci et al., 2011a, Pasqualucci et al., 2014). Alterations promote GC proliferation and block differentiation through their effects on tumour suppressor genes regulating B cell development. Normal KMT2D targets the following lymphoid tumour suppressor genes; *TNFAIP3, SOCS3, and TNFRSF14* (Molavi et al., 2013, Cheung et al., 2010b, Compagno et al., 2009). However following its inactivation, loss of methylation at lysine 4 on histone 3 (H3K4) in the enhancers and promoters of these genes, likely lead to transcriptional repression and deregulated GC B cell proliferation. CREBBP acetylates two non-histone proteins tp53 and bcl6, leading to activation of tp53 mediated apoptosis and inactivation of bcl6 function (Pasqualucci et al., 2011a). Therefore, loss of function results in inhibition of TP53 mediated apoptosis and enhanced bcl6 activity. Bcl6 is a transcription repressor with its expression carefully modulated during B cell development. It is highly expressed in GC B cells and prevents premature activation and differentiation of GC B cells while also allowing a tolerance to DNA breaks required for SHM and CSR of the immunoglobulin genes (Basso and Dalla-Favera, 2012). The enhanced activity of BCL6 may in part be responsible for GC expansion and maturation block. BCL6 is also recurrently altered by chromosomal

translocations and somatic hypermutation in FL (Akasaka et al., 2003, Bastard et al., 1994, Otsuki et al., 1995), with a higher prevalence of translocation frequencies found in patients who go on to transform to aggressive DLBCL (Akasaka et al., 2003).

EZH2, a methyltransferase that catalyses methylation at histone 3 lysine 27, is mutated in up to 30% of FL cases (Sneeringer et al., 2010). Mutations result in a gain of function, and as EZH2 is a transcriptional repressor, mutations lead to a permanent silencing of genes involved in cell cycle checkpoints and B cell differentiation, enabling rapid B cell proliferation within the GC. Individually, *KMT2D*, *CREBBP* and *EZH2* mutations were shown to promote lymphomagenesis in presence of BCL2 overexpression, as evidenced in the VavP-BCL2 mouse model (Zhang et al., 2015, Zhang et al., 2017, Beguelin et al., 2013, Berg et al., 2014).

The reversible nature of epigenetic modifications makes them an attractive drug target. To date, the FDA had licenced several epigenetic therapies in the treatment of haematological malignancies, including myelodysplastic syndrome and multiple myeloma (MM) (Garcia-Manero and Fenaux, 2011, Laubach et al., 2015). However, in FL, several trials using epigenetic modifying agents including histone deacetylase (HDAC) and *EZH2* inhibitors as single agent therapies have shown disappointing complete response rates (Evens et al., 2016, Morschhauser et al., 2017, Kirschbaum et al., 2011, Ogura et al., 2014).

1.12.3 Genetics of transformed FL

As previously mentioned, tFL is clonally related to the previous FL disease, sharing an identical t14;18 translocation. Therefore, the genomic profile of tFL consists of alterations both shared and distinct from the paired FL, which was illustrated in studies using paired FL-tFL samples. Recurrently shared mutations, occur in the following genes; *CREBBP*, *KMT2D*, *EZH2*, *TNFRSF14* and *FAS* (Okosun et al., 2014, Pasqualucci et al., 2014), indicating histone modification, apoptosis resistance and immune modulation as pathways affected in the CPC. Recurrent mutations specifically gained at transformation include *CDKN2A* deletions and/or loss of heterozygosity (Pasqualucci et al., 2014),

deletions and loss of heterozygosity affecting *TP53* and deregulation of *MYC* through point mutations, chromosomal translocations or copy number amplification (Lo Coco et al., 1993, Sander et al., 1993, Yano et al., 1992, Lossos et al., 2002). As these aberrations impact on cell proliferation and DNA damage pathways, it is likely that these processes contribute to transformation. Deletions in genes involved in immune recognition control such as *B2M* and *CD58* were also found at higher frequencies in tFL, suggestive of immune surveillance escape being a contributor. Several gene mutations found in both FL and t-FL but which could not be assigned to either shared or phase-specific due to their variability, included *STAT6*, *FOXO1* and *CARD11* (Pasqualucci et al., 2014), suggesting a heterogeneous contribution of multiple signalling pathways to both FL and tFL pathogenesis. In comparison to FL, tFL has a more complex genetic profile with distinguishing alterations in genes such as *MYC* and *TP53* (Lossos and Gascoyne, 2011, Davies et al., 2007). As the majority of DLBCL-like tFL resembles *de novo* GCB, a comparison between thirty nine tFL and one hundred and two de novo GCB-DLBCL cases revealed similar genetic alterations in REL, EZH2, GNA13, and TNFRSF14 mutations (Morin et al., 2010, Morin et al., 2011, Pasqualucci et al., 2011b). Additionally, t-FL also has specific aberrations rarely seen in *de novo* DLBCL including loss of *CDKN2A/B*, *STAT6* mutations, *ARID1A* mutations and *FAS* mutations/deletions (Pasqualucci et al., 2014). It is unlikely that a single driving oncogenic event exists for transformation due to the lack of uniformity across cases (Pasqualucci et al., 2014, Bouska et al., 2014, Davies et al., 2007) and instead requires a specific sequence of genetic 'hits' to confer an aggressive phenotype.

1.12.4 Genetic heterogeneity in FL

Recurrent genetic and epigenetic mutations in patients give us an indication into key pathways perturbed in the disease. Yet there is significant genetic heterogeneity in disease, reflecting the complex biology of disease. This has complicated identifying further hallmark mutations of disease, in addition to the t14;18 translocation. Furthermore, heterogeneity is not only restricted between patients but also in different subclonal populations within an individual. Araf *et al* performed profiling on spatially

separated, synchronous biopsies and found spatial discordance in several genes including *TNFAIP3*, *TNFRSF14* and *EP300* (Araf et al., 2018). This was in part due to differences in clonal dominance between sites, in which a mutation in a subclonal population at one site becomes clonally dominant at another site. The intra-tumour heterogeneity highlights the shortfalls of genetic landscape profiling in FL from single site biopsies and compromises the precision medicine approach so sought after in FL treatment, with Darwinian selection most likely to result in treatment failure. However, *CREBBP* and *KMT2D* mutations were found to be clonally maintained in spatially separated biopsies, reaffirming them as early events in pathogenesis and therefore a more robust target for therapeutic intervention.

1.13 Microenvironment

In addition to intrinsic genetic aberrations, cross-talk between lymphoma cells and their microenvironment is considered critical to tumour pathogenesis and survival. Similar to their normal counterparts, FL cells are admixed with immune cells, including stromal cells, macrophages and follicular helper CD4+ T cells in invaded follicles (Carbone et al., 2009). The difficulty of growing FL cells in vitro without stimulation through CD40 engagement (Ghia et al., 1998), a main signalling pathway of B and T cell interactions, supports a direct tumour-immune cell interaction. A pivotal study by Dave *et al* in which the gene expression of non-malignant immune cells was analysed elucidated how important the immune system, irrespective of tumour genetics, is in determining the clinical outcome (Dave et al., 2004). Enrichment for genes expressed in T cells was found to be associated with a favourable prognosis, whereas genes expressed in macrophages and follicular dendritic cells are strongly associated with poor clinical prognosis. However, results from subsequent immunohistochemistry (IHC) studies investigating how the number and arrangement of immune cells impacts on prognosis has been highly variable due to the heterogeneity of treatment and patient selection criteria (Carreras et al., 2009, de Jong et al., 2009, Lee et al., 2006).

In general, the germinal centre is a hostile microenvironment for B cells, in which those displaying low affinity BCRs for their cognate antigens are deprived of survival signal and

undergo apoptosis. One way in which FL cells subvert this Darwinian selection process is through constitutive expression of the anti-apoptotic protein, bcl2. However, there are additional mechanisms which support tumour survival. FL cells expressing high levels of CXCR4 and CXCR5 are recruited into follicles by chemokines secreted from follicular helper T cells (T_{FH}) and dendritic cells, including CXCL13 and CXCL12/SDF-1 (Lopez-Giral et al., 2004, Ame-Thomas et al., 2007, Husson et al., 2002). FL cells also promote the secretion of CCL2 from stromal cells which although having no direct effect on B cell survival, leads to monocyte recruitment and therefore modulation of the microenvironment (68). Mesenchymal stromal cells (MSC) were also found to drive monocyte differentiation towards a pro-angiogenic and anti-inflammatory tumour associated macrophage (TAM)-like phenotype, with up-regulation of IL-10, IL-6 and VEGF (Guilloton et al., 2012).

IL-4 is up-regulated within the FL microenvironment and is predominantly produced by T_{FH} cells (Pangault et al., 2010). IL-4 binds to its receptor on tumour cells and activates STAT6 responsive genes that may promote survival of the tumour cells (Lu et al., 2005). It also stimulates the production of CCL17 (Rawal et al., 2013), which facilitate recruitment of regulatory T cells (T_{Regs}) that may generate an immunosuppressive tumour microenvironment and promote immune evasion (Zou, 2005). In addition, the gene expression profile of tumour infiltrating lymphocytes were also shown to have an altered gene expression profile that corresponds to impaired actin polymerization and immune synapse formation, leading to T-cell exhaustion that can be repaired with the immunomodulatory drug lenalidomide (Ramsay and Gribben, 2011, Ramsay et al., 2009).

Therefore, the FL cell niche both promotes tumour growth and organises the microenvironment through cell recruitment, polarisation and survival signalling to provide a pro-tumour environment. Mechanisms for how this is achieved remains to be fully elucidated.

1.13.1 N-linked glycosylation of follicular lymphoma surface immunoglobulin

N-linked glycosylation (N-gly) is a common protein modification that occurs co-translationally within the endoplasmic reticulum (ER) of eukaryotic cells (Aebi, 2013). A fourteen sugar precursor is covalently bound to the carbamino nitrogen on asparagine residues within the specific glycosylation amino acid motif N-X-S/T (where N stands for asparagine, X could be any amino acid except proline, S stands for serine and T stands for threonine) on the nascent protein. The glycan then undergoes trimming and maturation in the Golgi to transform from a high mannose oligosaccharide to a complex glycan, as summarised in Figure 1.7. Complex glycoproteins can be transported to the cell surface, where they can take part in cell–cell interactions required for adhesion, migration and signalling (Varki, 1993, Saxon and Bertozzi, 2001).

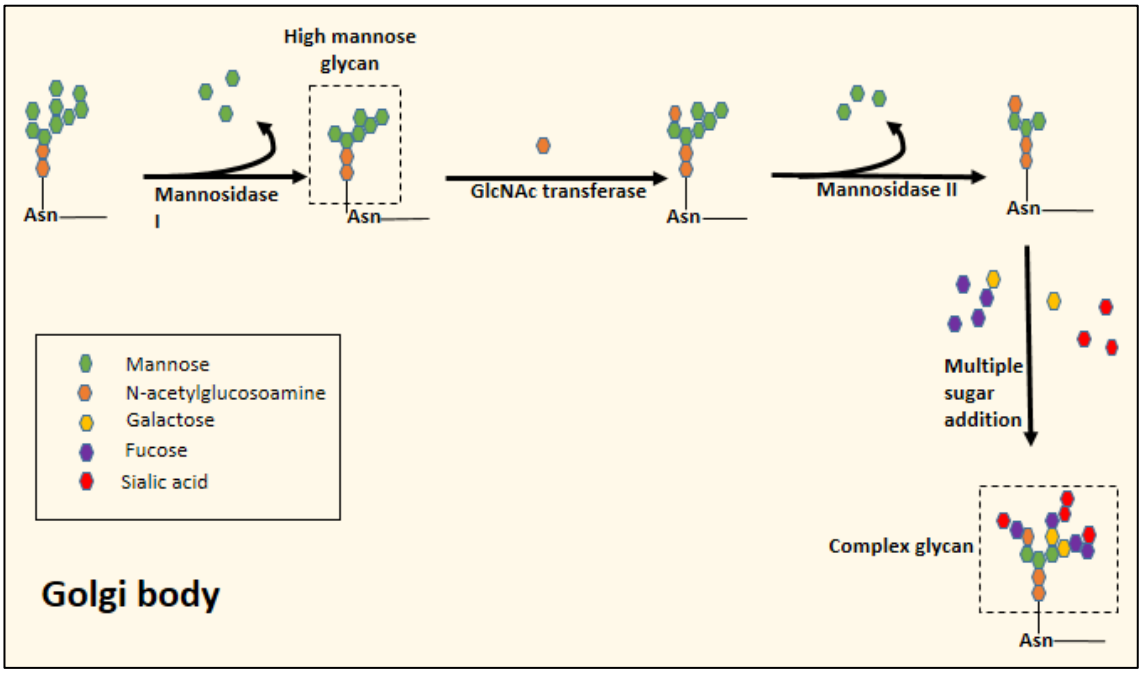


Figure 1.7: *N-glycosylation trimming and maturation in the Golgi.* The fourteen sugar precursor glycan attached to the asparagine (Asn) reside of the N-X-S/T amino acid motif undergoes trimming in the ER through the action of glycosidases. This results in the removal of 3 terminal glucose residues (not shown here). The glycoprotein is then trafficked to the Golgi where Mannosidase I removes terminal mannose residues, leaving a high mannose glycan. N-acetylglucosamine transferase (GlcNAc transferase) transfers the monosaccharide N- acetylglucosamine to the glycan chain, followed by more trimming and addition of different sugar molecules. This results in the production of branched, complex oligosaccharides.

N-gly motifs are found within the immunoglobulin, mainly confined to the constant regions (Radcliffe et al., 2007). During B-cell development, an N-linked glycosylation site in the first constant region of the μ-chain is essential for a functional pre-B cell receptor (Ubelhart et al., 2010). A few germline encoded variable genes also carry sites, including *IGHV1–08*, *IGHV4–34* and *IGHV-5a* (Radcliffe et al., 2007). Interestingly, during the SHM process of FL Ig, the number of N-gly motifs in the variable region increases significantly, with the majority occurring in the heavy chain compared to the light chain (Zhu et al., 2002, Zabalegui et al., 2004). Motifs are most prevalent within the CDR regions (McCann et al., 2008), which are associated with antigen binding. The acquisition of at least one motif occurs in >90% of FL cases (McCann et al., 2008). Interestingly, this phenomenon also occurs at high frequencies in other GC centre B cell lymphomas including DLBCL and Burkitt's lymphoma (Zhu et al., 2002). Acquisition in non-GC B cell lymphoproliferative disorders, including multiple myeloma and CLL, is considerably lower, with an incidence range of 3-18%, similar to the value seen in normal somatically mutated B cells (~9%) (Alcoceba et al., 2012). Importantly, acquired sites in FL are conserved during ongoing mutation, in contrast to germline-derived sites (McCann et al., 2008).

To determine the glycosylation profile at these introduced sites, IgM proteins derived from the B cells of FL patients were exposed to specific glycosidases. Analysis revealed high mannose-terminating glycans (Coelho et al., 2010) (Figure 1.8), an unusual finding as surface proteins normally display branched and complex oligosaccharides that have undergone processing in the ER and Golgi (Aebi, 2013, Varki, 1993, Saxon and Bertozzi, 2001). Interestingly, N-glycosylation sites in the constant region are fully processed (Radcliffe et al., 2007), suggesting that normal Golgi transit is maintained and sugar processing enzymes remain functional. This infers that the added oligomannose have a tumour related function. This is supported by the observation that all somatically mutated N-gly sites in FL are oligomannosylated (McCann et al., 2008). Interestingly, germline-encoded N-gly sites such as that found in the CDR2 region of *V4-34* are not glycosylated (McCann et al., 2008), suggesting introduced sites have a biological significance in FL.

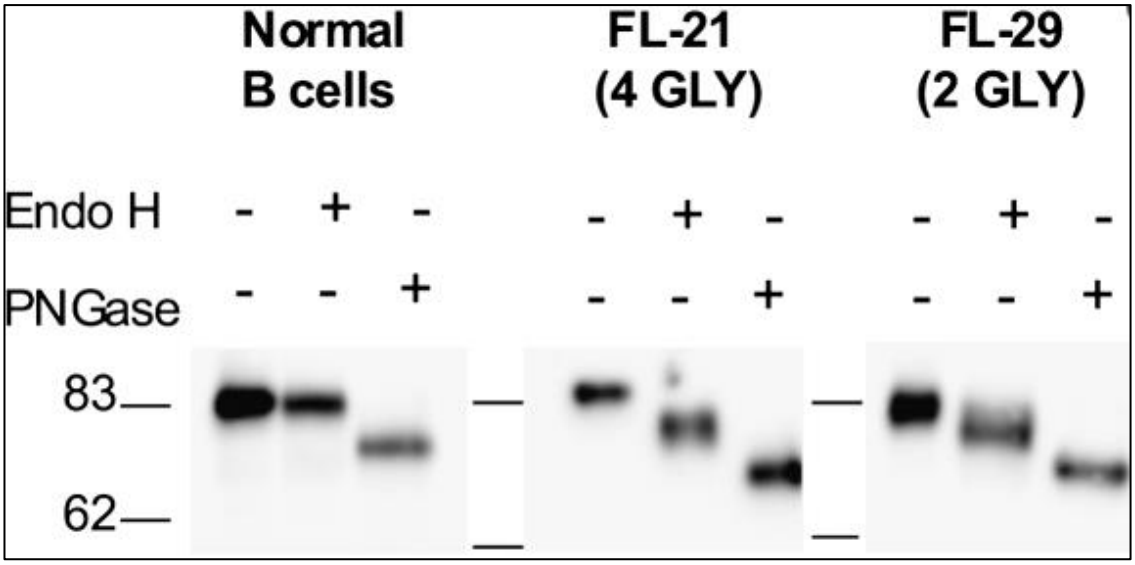


Figure 1.8: *Mannosylation of sIg derived from two FL patients with known N-gly sites*. Endoglycosidase H (Endo-H) and Peptide:N-Glycosidase F (PNGase) cleave either mannosylated or all *N*-glycans, respectively. FL proteins contain additional mannosylated glycans as seen through increased mobility of protein following Endo-H treatment. The distance of migration correlates to the number of N-gly sites in the heavy chain variable region (indicated in brackets), suggesting all sites are mannosylated. Image is taken from Coelho *et al* (Coelho et al., 2010).

N-glycans were found to severely impair cognate antigen binding in the CDRs and antigen-induced BCR stimulation (Radcliffe et al., 2007, Schneider et al., 2015). This was indirectly supported by Sabouri *et al* discovering how N-glycans block self-antigen binding on anergic B cells, modulating antibody avidity and redeeming autoantibodies from self-reactivity (Sabouri et al., 2014). The prevalence of sites within the antigen binding CDRs, suggests that oligomannoses interact with their microenvironment through an antigen-independent mechanism, enabling FL residence and proliferation within the GC.

1.13.2 Interaction between oligomannosylated Ig and C-type lectins

Binding studies using oligomannosylated FL sIg and mannose-binding lectin (MBL), confirmed that glycans were accessible for binding to calcium-dependent (C-type) lectins (McCann et al., 2008), suggesting a novel interaction between tumour cells and

their microenvironment through oligomannose engagement. C-type lectins are pattern recognition receptors (PRR) that bind to glycans through their carbohydrate binding domains (2009). They have a wide variety of roles including phagocytosis/antigen presentation of pathogens and mediation of endogenous cell-cell interaction during the immune response (Dambuza and Brown, 2015, Geijtenbeek and Gringhuis, 2009). As genes associated with macrophages and dendritic cells are predictive of poor prognosis (Dave et al., 2004), it was hypothesised that lectins expressed on these immune cells were candidate molecules for interaction with sIg mannoses. Two major C-type lectins were proposed which have specificity for high mannose structures and are expressed by macrophages and dendritic cells; mannose receptor (MR) and DC-SIGN (DC-specific intercellular adhesion molecule-3–grabbing non-integrin). In vitro studies revealed that mannosylated sIgM from primary FL cells interacted with both lectins and induced sIgM-associated intracellular calcium fluxes, highlighting a functional interaction (Coelho et al., 2010). In addition, recombinant human DC-SIGN incubation with IgM positive FL cells led to phosphorylation of BCR signalling kinases, including ERK1/2 and AKT, and increased expression of MYC (Linley et al., 2015, Amin et al., 2015). This highlights the activation of the BCR pathway (and its associated targets) through an antigen-independent mechanism. Although phosphorylation of kinases was delayed and reduced in comparison to control anti-IgM treatment, the duration of response was longer (Linley et al., 2015). Furthermore, the sIg failed to undergo endocytosis, suggesting that constitutive activation of the BCR signalling pathway through DC-SIGN engagement is enough to maintain FL survival. These effects were lost with the treatment of Btk and Syk inhibitors, providing support to BCR activation via the lectin-mannose interaction (Linley et al., 2015). The proposed mechanism is summarised in Figure 1.9. DC-SIGN was found to be strongly expressed on M2 macrophages and dendritic cells derived from FL samples (Amin et al., 2015) which may have relevance to the poor prognosis associated with a high CD68+ macrophage count (Dave et al., 2004). It was also found to be in direct contact with FL cells *in situ*, indicating the feasibility of DC-SIGN as a FL specific BCR ligand within the disease niche (Amin et al., 2015). Bacterial lectins *Pseudomonas aeruginosa* and *Burkholderia cenocepacia* were also shown to bind to FL sIg and induce sIg-associated intracellular calcium fluxes (Schneider et al., 2015), indicating the vast array of potential lectin candidates that may be utilised in disease

pathogenesis. There is a dispute regarding the Ig isotype able to bind DC-SIGN. Linley *et al* found both IgM+ and IgG+ FL cell were able to bind to DC-SIGN and activate signalling (Linley et al., 2015). However, Amin *et al* found that IgG+ FL cells were poorly activated by DC-SIGN (Amin et al., 2015). These discrepancies may be due to technical and IgG+/IgM+ case variation between the two studies. However, as IgG+ FLs are more commonly self-reactive (Sachen et al., 2012), the lowered affinity to lectins may be due to alternative stimulation (via autoantigen) of the BCR within this isotype subgroup. This warrants further investigation.

While DC-SIGN and MR are well described in innate immunity (Engering et al., 1997, Geijtenbeek et al., 2000), the evidence is accumulating regarding their role in tumour immunity. Cross-linking of MR on immature dendritic cells with anti-MR antibody was shown to activate DC maturation and induce the secretion of anti-inflammatory cytokines and Th2 attracting chemokines, leading to negative regulation of Th1 polarised responses (Chieppa et al., 2003). Interestingly, similar responses are seen when the receptor is engaged with cancer-derived glycoproteins. For example, ligation of MR with the ovarian cancer derived glycoprotein TAG-72 on TAMs lead to increased IL-10 production and decreased CCL3 (Allavena et al., 2010), the Th1 attracting chemokine. This immune dampening response is also seen when DC-SIGN is engaged with colon cancer derived glycoproteins, leading to an increased IL-10 expression (Nonaka et al., 2008). IL-10 production promotes polarisation of M2 macrophages (Allavena et al., 2010), the phenotype displayed by FL derived TAMs. MR and DC-SIGN are highly expressed on TAMs (Amin et al., 2015), suggesting that these lectins play a role in polarising the immune cells to create a pro tumour, immune suppressive microenvironment. This suggests another dimension to the lectin-mannose interaction in the support of tumourigenesis.

Mannose-lectin interactions motifs could enable FL cells to survive and accumulate mutations within the hostile GC through both stimulating the BCR and creating a tumour supportive microenvironment. Determining how N-gly motifs influence disease evolution could provide validation for targeting this tumour-microenvironment interaction. Promising evidence has been seen in a case study in which ISFN cells gathered from several lymph nodes of a 69 year old male contained a conserved N-gly

motif (NCS) in the CDR3 region of the *IGHV*, suggesting N-gly motifs can be acquired within malignant precursor cells, perhaps in parallel with the t(14:18) translocation (Kosmidis et al., 2017). This was supported in another case study of a 35 year old male in which clonally related ISFN and FL cells shared an acquired motif within the CDR2 region of the *IGHV* (Mamessier et al., 2015). Having a functional mannose-lectin interaction at this early stage in disease development may explain how precursor cells are able to survive in the hostile germinal centre without the need for high affinity BCRs enabling accumulation of genetic hits required for malignant transformation.

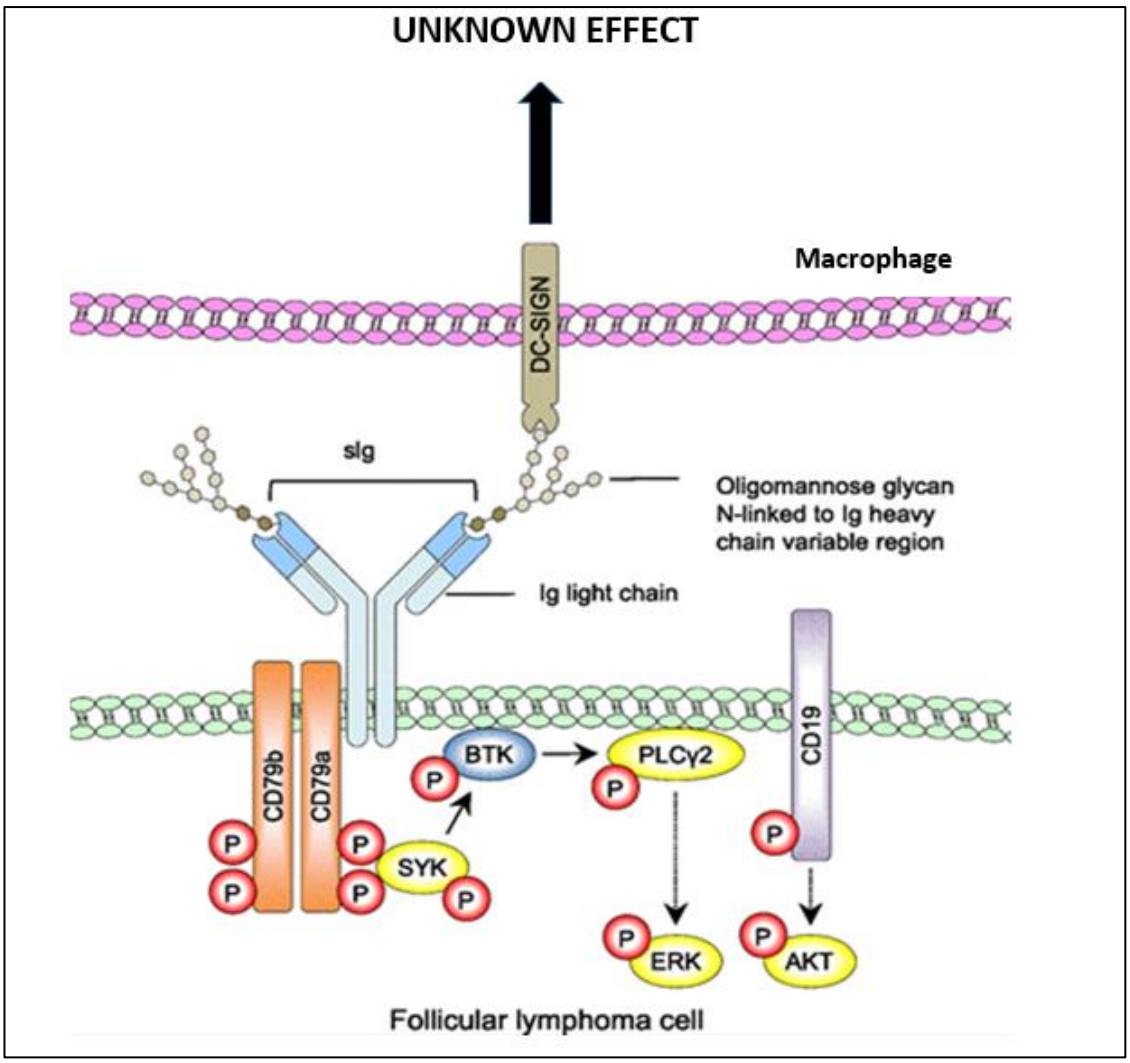


Figure 1.9: *Model highlighting the inferred interaction between FL cells and the microenvironment via the mannose-lectin interaction*. N-glycosylation motifs located in the variable region of the sIg on FL cells is an acceptor site for high mannose glycans. C type lectins including DC-SIGN are expressed on macrophages and dendritic cells. These selectively bind to the high mannose structures, leading to the organisation of the BCR and CD19 co-receptor signalling platforms and downstream phosphorylation of signalling kinases, including SYK, BTK, ERK and AKT (10). How this interaction may affect the partnering immune cell expressing DC-SIGN/MR in the context of FL has not been investigated. Image adapted from Strout, 2015 (Strout, 2015).

1.14 Current challenges in FL

Despite improvements to patient survival following the introduction of Rituximab, FL remains an incurable disease with a relapse-remitting pattern. The varying clinical

outcomes of patients, with some having durable responses to first-line treatment whilst others progress or transform suggest an underlying intrinsic biological heterogeneity in the disease. This has been exemplified by NGS that has revealed a complex and variable genetic landscape. Recurrent mutations in specific genes have been identified, indicating biological pathways perturbed in disease, including epigenetic regulation. However, these mutations are not found within all patients and vary in clonal dominance, not just between patients but also in spatially separated biopsies from individual cases (Araf et al., 2018). Heterogeneity is also illustrated by the vast number of minor subclones identified from *IGHV* sequences (Carlotti et al., 2015). Targeting a mutation in a clonal population may therefore enable a minor subclone to gain dominance, leading to relapse. This heterogeneity of subclones and their genetics compromises the approach of precision medicine. Targeting early events in pathogenesis which are conserved clonally would likely have the most beneficial impact on sustained remission, which is why characterising events of the putative CPC compartment is an important area of research in FL. Genetic changes in the CPC are being investigated and characterised (Okosun et al., 2014), yet CPC-microenvironment interactions remain unexplored despite our extensive knowledge on the importance of microenvironment-tumour cell interactions in pathogenesis.

1.15 Aims of this thesis

Acquired N-gly motifs in the variable region of the *IGH* are believed to possess a pathogenic function due to their high frequency in FL cases and rarity in healthy somatically mutated B cells. This is likely conferred by their attachment to high mannoses and subsequent interaction with lectins, which activate the BCR signalling pathway. Determining at what stage of tumour evolution these motifs are acquired will validate their role in disease development.

The primary aim of this thesis was to determine whether N-gly motifs are acquired within the putative CPC population by investigating the behaviour of N-gly sites during ongoing SHM and disease progression. To do this we analysed the incidence and

maintenance of sites within the heterogeneous clonal repertoire of several patients taken at different time points of disease, ranging from diagnosis to transformation. Temporal samples were either taken from the same or different anatomical sites. Clones were derived through targeted sequencing of the *IGHV* and subclonal relationships were visualised through lineage trees. To our knowledge, this is the first study that has analysed the relationship between FL progression and N-gly sites in a variable patient cohort who have undergone different lines of therapy and presented with different clinical courses, reflecting the heterogeneous nature of the disease.

The second aim was to use a unique, in-depth multiplex phenotyping method to investigate the tumour microenvironment with regards to expression of the C-type lectins, DC-SIGN and MR. Using digital image analysis software, the *in situ* frequency, distribution and spatial arrangement of lectin-expressing immune cells was analysed in tissue sections derived from a heterogeneous FL patient cohort. Patients had an overall survival of less than five years or greater than fifteen years, enabling us to compare the microenvironmental landscape in extreme survival cases and investigate the prognostic value of lectin expression in FL disease.

Chapter 2: MATERIALS AND METHODS

2.1 Patient samples and ethics

For sequencing studies, readily available genomic DNA samples were previously extracted from lymph node biopsies of patients affected with FL at different time points of disease including diagnosis, relapse and tFL. Written consent was obtained for collection and use of specimens for research purposes with ethical approval obtained from the Institutional Review Board under the following submissions: 10/H0704/65 and 06/Q0605/69. Clinical information was available through the database maintained within the Centre for Haemato-Oncology at Barts Cancer Institute.

For immunohistochemistry studies, whole section slides were prepared from diagnostic paraffin-embedded lymph node blocks from patients diagnosed with FL at St Bartholomew's Hospital (London, United Kingdom). Ten slides were prepared for patients whose survival was less than five years and ten slides were prepared for patients whose survival was greater than fifteen years following diagnosis. Whole sections from reactive lymph nodes of ten healthy subjects were used as a control. Ethical approval for this study was obtained from East London and the City HA Local Research Ethics Committee under the submission 05/Q0605/140.

All tumour biopsies were reviewed by a panel of haemato-pathologists to confirm the diagnosis. Further description of samples can be found within the relevant results chapters of this thesis.

2.2 Cell lines

Germinal centre derived B-NHL cell lines (SU-DHL4 and SU-DHL6) were previously obtained from Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ). Both cell lines were t(14:18) positive. SU-DHL4 expresses IgG whereas SU-DHL4 expresses IgM, the two immunoglobulins dominantly expressed on FL cells. Cryovials

containing cells passaged four times (stored in 10% DSMO, 10% foetal calf serum (FCS) and RPMI 1640 medium) were removed from liquid nitrogen storage (-196^{o}C) and thawed in a water bath set to 37^{o}C. 1ml of 37^{o}C RPMI 1640 medium supplemented with 1% penicillin/streptomycin and 10% FCS were added to the vial and transferred to 15ml falcon tubes containing 9ml of medium before spinning at 1000rpm for 5 minutes at room temperature (20^{o}C). Supernatant was discarded and the cell pellet was resuspended in 10ml of medium before transfer to a sterile T75 tissue culture flask and incubated in a humidified 5% CO2 incubator at 37^{o}C. Following an initial slow cell growth due to acclimatizing to new conditions, cells were passaged approximately every three days to maintain optimal cell density and viability as indicated from cell line profiles on the DSMZ website. Morphology and density of cells was assessed under light microscope as well as to check for contamination signs. Mycoplasma testing was carried out regularly.

Cell viability and density measurements were carried out using the LUNA-II™ Automated Cell Counter (Logos Biosystems). Measured viabilities were determined by a dye exclusion method, using Trypan blue staining. Live cells do not take up the dye due to their intact cell membrane leaving the cytoplasm unstained. Therefore, cells stained blue are considered dead. Following thorough resuspension of cells, 10µl of cell sample was mixed with 10µl of Trypan blue dye before being loaded into a cell-counting chamber on the Luna slide. Similar to normal haematocytometers, the Luna performs multiple cell counts in a known area and outputs a mean value. >85% viability highlighted adequate growth conditions, in terms of cell density and nutrient supply.

2.3 DNA extraction from cell lines

Total DNA extraction was carried out using the DNeasy Blood and Tissue kit (Qiagen, Cat No. 69504). 5 x 10^6 cells were taken from the cell suspension, and centrifuged for 5 minutes at 300 x g at room temperature. Supernatant was discarded and cell pellet resuspended in 200µl of phosphate buffered saline (PBS). 20µl of proteinase K was added to the pellet along with 200µl of Buffer AL. Sample was pulse vortexed briefly and mixed to ensure effective lysis. Proteinase K is a serine protease that eliminates proteins

and DNases that may affect DNA. Buffer AL contains a chaotropic salt, which has an additive effect to protein destabilisation by disrupting hydrogen bonds in water, which denatures biomacromolecules. Furthermore, chaotropic salts disrupt the interaction of DNA with water and enable it to bind to silica. Samples were incubated at 56^{o}C for 10 minutes. 200µl of 100% ethanol was added to samples and pulse vortexed. Ethanol precipitates nucleic acids out of solution. The mixture was transferred to a QIAmp Mini Spin Column placed inside a 2ml Eppendorf. The column was centrifuged at 6000 x g for one minute. Precipitation by ethanol enables the separation of nucleic acids from the rest of the solution with the DNA binding to the silica membrane and contaminants passing through. 500µl of Buffer AW1 was added followed by a centrifuge at 6000 x g for 1 minute. Flow through was discarded. 500µl of buffer AW2 was added and sample was centrifuged at 20000 x g for 3 minutes. These washing steps are essential for removing residual contaminants from the DNA including chaotropic salts. A dry spin of two minutes ensures removal of ethanol. Finally, 200µl of Buffer AE is added to the spin column followed by a five minute stand period. The column is centrifuged at 6000 x g for 1 minute, resulting in the elution of DNA from the silica into a 1.5ml eppendorf. Purified DNA was stored at 4^{o}C.

2.4 IGHV amplification

We sought to amplify the *IGHV* from FL samples. Due to the vast number of V, D and J segments, we used the widely validated Biomed consortium's multiplex PCR approach in which all functional VH family members were able to be recognised by one of 6 VH consensus primers, targeting the Framework 1 region (van Dongen et al., 2003). This is due to 6-7 oligonucleotides in each primer binding to its corresponding VH family member with no mismatch. These primers were used in conjunction with a single JH consensus primer designed to anneal to the 3' end of all six known JH gene segments (Figure 2.1). Taq polymerase lacks a proofreading 3'-5' exonuclease domain and therefore cannot remove a misincorporated nucleotide. As single nucleotide polymorphism (SNP) analysis is critical to identifying acquired N-glycosylation motifs and analysing SHM-based evolutionary pathways, minimising SNPs generated by polymerase

reading errors required using a high fidelity polymerase. This was achieved by using the Phusion Green Hot Start II High-Fidelity PCR Master Mix, comprising of the Phusion Hot Start II High-Fidelity DNA Polymerase, nucleotides, a reaction buffer containing $MgCl_2$, a density reagent that permits direct loading of the PCR products to the gel, and two tracking dyes. The tracking dyes were bromophenol blue and xylene cyanol FF that comigrates with approximately 300bp fragments and approximately 4000bp fragments, respectively, enabling the rate of migration of DNA molecules to be monitored. The master mix reduces the amount of pipetting steps and therefore lessens the risk of contamination during the reaction set up. A 10µl reaction consisting of 5µl of master mix, 0.5µl of FR1 primer mix (each individual primer at 10µM, 0.5µl of JH primer (10µM), 2µl of nuclease-free water and 2µl of genomic DNA (20ng/µl). The polymerase activity is inhibited at ambient temperature and therefore reaction setup can be done at room temperature, without the risk of primer dimer formation and non-specific amplification products. A no template control (NTC) consisting of all components apart from DNA was included in all experimental runs to exclude contamination with exogenous DNA. The cycling conditions consisted of an initial denaturation step at 95°C for 30 seconds, followed by 40 cycles of DNA denaturation at 95°C for 30 seconds, then primer annealing at 58°C for 30 seconds, then primer extension at 72°C for 30 seconds and a final extension step of 72°C for 7 minutes. Following heteroduplex analysis and Sanger sequencing of the homoduplex band (described below), the PCR was repeated using a single FR1 primer corresponding to the IMGT-VQUEST identified VH family.

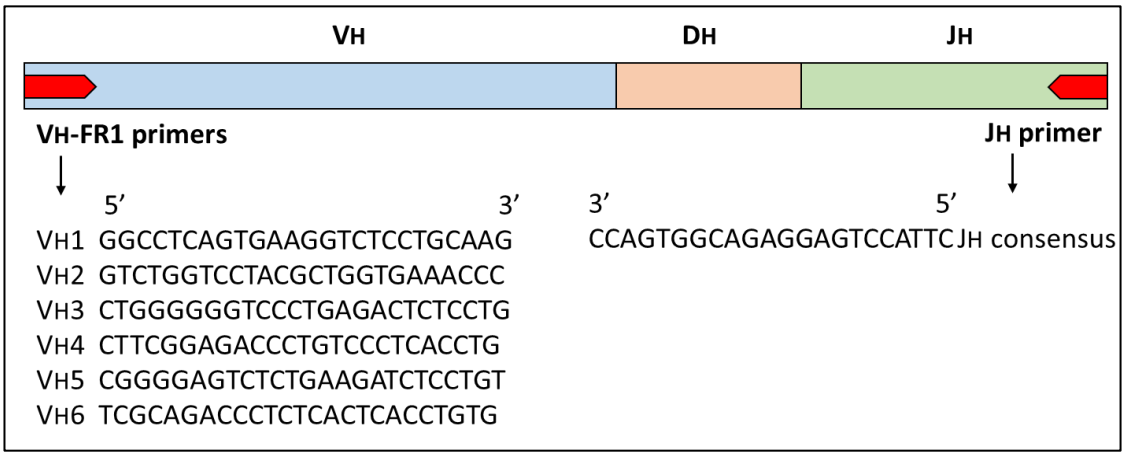


Figure 2.1: *Molecular organisation of an IGH VH-JH rearranged sequence in genomic DNA*. VH family FR1 primers and a single consensus JH primer were used to target and amplify this region during multiplex PCR. Primer sequences (indicated) were taken from the BIOMED2 protocol (van Dongen et al., 2003).

2.5 Heteroduplex analysis

A drawback of the PCR technique in lymphoma-based *IGHV* studies is the risk of false positive results generated by amplification of *IGHV* from polyclonal B cells, especially when there is a high background of polyclonal B cells existing in the tumour sample. To discriminate between monoclonal and polyclonal PCR products, heteroduplex analysis is used. In heteroduplex analysis, PCR products are denatured at high temperatures and then reannealed by rapid cooling to create duplexes. Homoduplexes are formed by monoclonal PCR products (with identical junction regions) whereas heteroduplexes are formed by polyclonal PCR products (with heterogeneous junction regions), resulting in a smear of slow migrating fragments on a polyacrylamide gel (Figure 2.2).

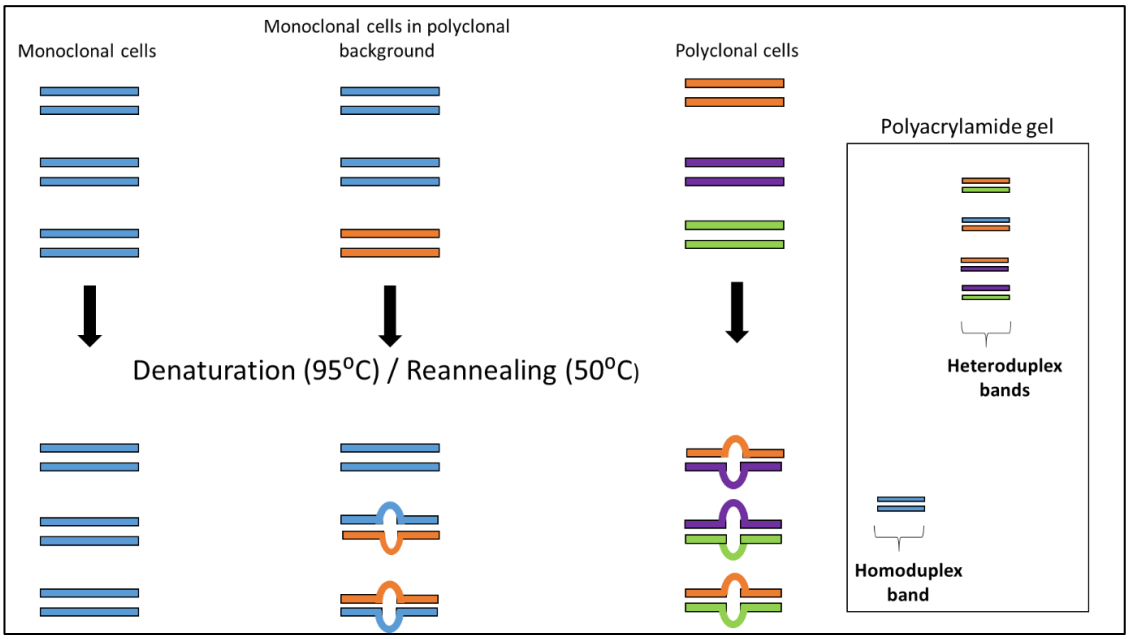


Figure 2.2: *Schematic diagram of the heteroduplex analysis technique*. PCR products are heat denatured and quickly cooled to induce duplex formation. In cell samples consisting of clonal B cells (left panel), PCR products give rise to homoduplexes whereas samples with clonal and polyclonal B cells present give rise to homo- and heteroduplexes (middle panel). Polyclonal samples will mainly form heteroduplexes (right panel). Homoduplexes have matched junctional regions and therefore migrate more rapidly through the polyacrylamide gel. Heteroduplexes have mismatched junctional regions and therefore have an open conformation at these sites, causing them to migrate slowly through the gel, enabling distinction between monoclonal and polyclonal samples.

The clonal gene rearrangements in lymphoma cells were distinguished from polyclonal normal cells using heteroduplex analysis. PCR products were incubated at 95^{o}C for denaturation and then rapidly cooled to 50^{o}C for 1 hour for reannealing. Samples were then promptly put on ice and 1x gel loading dye was added. Samples were added to wells of non-denaturing 10% polyacrylamide gel (Novex) with one additional well containing a 100bp DNA ladder as a size reference. Electrophoresis was performed with 20mA applied to the gel in 1x TBE running buffer for 14 hours at 4^{o}C. The gel was then submerged in 1x TBE buffer containing 3x GelRed (Biotium) and gently agitated at room temperature for 30 minutes. Bands were then viewed under ultraviolet (UV) light. Prominent homoduplex bands located at the 300-400bp mark represented the sequence of the *IGHV* region of the major clone (highest count number) whereas faint bands located nearer to the top of the gel represent non-malignant polyclonal PCR products (Figure 2.2). Homo-duplex bands were extracted, and DNA recovered using QIAquick® Gel Extraction Kit, followed by re-amplification by PCR. Following visualisation by agarose gel and UV light, PCR products underwent clean up and Sanger sequencing.

2.6 Gel extraction

Bands located at the 300-400bp mark which corresponds to the amplified *IGHV* were excised from the gel. DNA was extracted using the QIAquick® Gel Extraction Kit (Qiagen) following the manufacturer's instructions.

2.7 Agarose gel electrophoresis

IGHV PCR products were visualised on agarose gels. 2% agarose gels were prepared using 2g of agarose in 100ml of 1x TBE buffer. Agarose was melted in the buffer by heating. 5µl of GelRed (Biotium) was promptly added to the gel. GelRed is a fluorophore that intercalates with nucleic acids resulting in DNA visualisation under UV light. The gel mix was poured into an appropriate sized gel cast tray and combs applied to create wells

for sample loading. Once the gel was set following a twenty minute period, the gel was placed in 1x TBE running buffer. 15µl of PCR product was added to each well with an additional well containing a 100bp DNA ladder (exACTGene™, Fisher Scientific) used to identify amplicons of the appropriate size. Electrophoresis was performed for one hour at 100V. As DNA is negatively charged, it migrates towards the positive electrode through the agarose pores with small fragments migrating faster and large fragments migrating slower. DNA bands were visualised by transillumination under UV light that excites the fluorophore, producing a bright orange colour. Images were collected digitally.

2.8 PCR product clean-up

To remove any unincorporated nucleotides and primers, enzymatic clean up preceded sequencing to minimise interference. The ExoSAP-IT kit (Thermofisher) was used following manufacturers guidelines. The minimal pipetting steps reduce contamination of the product. 2µl of ExoSAP-IT reagent was added to 5µl of PCR sample, followed by an incubation step of fifteen minutes at 37^{o}C. This activates two enzymes, Exonuclease I and Shrimp alkaline phosphatase in the ExoSAP-IT reagent. Shrimp alkaline phosphatase removes 5' phosphates from unincorporated nucleotides whilst Exonuclease I specifically digests single-stranded DNA containing a 3' OH group in the 3' to 5' direction. This selectivity ensures that the double-stranded PCR product is not affected by these enzymes and only residual primers and unincorporated nucleotides are removed. An incubation step at 80^{o}C for fifteen minutes inactivates the enzymes. PCR products were stored at -20^{o}C or proceed directly to sequencing.

2.9 Nucleic acid assessment: Nanodrop

DNA quality and quantity were assessed by spectrophotometry using the NanoDrop ND-1000 Spectrophotometer (ThermoFisher). Nucleic acids absorb at a wavelength of 260nm. The amount of absorbance of a sample generates an optical density (OD) value

(OD260) which is multiplied by a factor of 50 to determine DNA concentration (ng/µl), which multiplied by the total volume, gives the overall DNA quantity of the sample. The TE buffer used to resuspend the DNA was used to create a baseline value to calculate DNA concentration. The OD value at 280nm is used to create the 260nm/280nm ratio which indicates the purity of DNA. Abnormalities in the 260/280nm ratio can indicate contamination of the sample with protein or reagents such as phenol which are used in DNA gel extraction methods. The OD value at 260nm is used to establish the 260nm/230nm ratio, a further assessment of purity. As many contaminants associated with the DNA extraction process absorb at 260nm and 230nm, a low ratio indicates a low DNA purity. A ratio of ~1.8 and ~2-2.2 for the 260nm/280nm and 260nm/230nm ratios, respectively, is indicative of contaminant-free DNA.

2.10 Nucleic acid assessment: Qubit

DNA destined for NGS platforms was quantitated using the Qubit® 2.0 Fluorometer (Invitrogen, Life Technologies), an approach that uses fluorescent dyes to selectively bind to either double or single-stranded DNA, RNA or protein, making it a more sensitive approach than UV absorbance by the Nanodrop. To quantify double-stranded DNA, the Qubit® dsDNA BR Assay Kit was employed. A range of standards was made up using 190µl of Qubit working solution and 10µl of the appropriate Qubit standard. Standards were used to calibrate the machine. 2µl of sample was added to 198µl of Qubit working solution and mixed by a pulse vortex. Samples and standards were incubated at room temperature for 2 minutes. Following calibration with the standards, samples were read on the machine and concentrations given in µg/µl. To gain the real concentration of the sample before its 2:200 dilution, the following equation was used;

$$\text{Concentration of sample} = \text{QF value} \times (200/2)$$

where: QF value = the value given by the Qubit® 2.0 Fluorometer

2 = the number of microliters of sample added to the assay tube.

2.11 Sanger sequencing

5µl of purified PCR product (~20ng/µl) mixed with 2.5µl of 10µM JH primer were sent to the Eurofins Genomics Sanger sequencing services (Germany) for bidirectional capillary electrophoresis using the ABI PRISM 3730XL Genetic Analyser. Sanger sequencing is based on synthesising DNA strands complementary to a single-stranded DNA template, therefore, a denaturation step of ~95^{o}C is required. In addition to the template DNA, it requires DNA polymerase, complementary DNA primer, the four deoxynucleotide triphosphates (dNTPs) and four modified nucleotides (ddNTPs), each labelled with a distinct fluorescent dye and lacking a 3' hydroxyl group. When ddNTPs are incorporated into the DNA strand, they prevent further extension by DNA polymerase as the absence of a hydroxyl group inhibits phosphodiester bond formation between two nucleotides. As a result, fragments of different sizes are produced which then undergo capillary electrophoresis. The exciting of the fluorophore generates a fluorescent signal, identifying the nucleotide in the original DNA template. Fragments of every possible length are generated through stoichiometric manipulation of the reaction components, revealing the entire reverse sequence of the template strand.

2.12 Analysis of sequencing results: IMGT/V-QUEST

AB1 and .seq files provided by Eurofins Genomics were analysed and interpreted by BioEdit Sequence Alignment Editor and Lasergene EditSeq. Electropherograms were checked by eye to ensure sequences were of good quality and base calls reliable for further analysis. This included ensuring peaks were evenly spaced and minimal baseline noise was present. The beginning and end sequences of the chromatogram were discarded due to lack of reliability. The FASTA format of sequences was submitted to IMGT/V-QUEST, a human Ig reference database that aligns nucleotide sequences to the V, D and J genes with the closest sequence homology (www.imgt.com). It also identifies the number and location of P- and N-nucleotides insertions at the junctional regions which occur during the somatic recombination process. As the tool can identify sequences in the reverse orientation and perform complementary reverse, no data

manipulation is required with the sequence generated by JH primer. Once the V gene was identified, the Sanger sequencing process was repeated with the JH primer being replaced by the appropriate VH Framework 1 primer. This was to accurately sequence the CDR3 region so that all variable region gene segments were identified, including the D and J segments. Furthermore, this ensured that identified mutations in the first Sanger sequencing were accurate.

2.13 Identification of acquired N-glycosylation sites

To determine the N-gly sites, the NetN-glyc 1.0 online server was used (www.cbs.dtu.dk/services/NetNGlyc/). Nucleic acid sequences were first translated using Transeq (EMBOSS) and entered in FASTA format. The server screens for consensus N-linked glycosylation motifs, consisting of asparagine (N), any amino acid except proline (X) and either serine or threonine (S/T). It also predicts which N-gly motifs are in fact glycosylated, a useful tool as we did not have access to primary FL cells from our paired FL-tFL samples to carry out glycosidase treatment of sIg (McCann et al., 2008, Coelho et al., 2010). To do this, nine neural networks interrogated each input sequence for N-X-S/T motifs, generating an average glycosylation potential score at the motif site. When a potential of >0.5 is reached and ≥7/9 networks are in agreement, the motif site is considered glycosylated. Motifs with proline as the middle amino acid are not considered glycosylated by the server, due to the amino acid causing steric hindrance of the asparagine residue (Gavel and von Heijne, 1990). To determine whether N-gly motifs were naturally occurring (germline-encoded) or acquired through SHM, sequences were aligned to germline immunoglobulin sequences in IMGT/V-QUEST.

2.14 Analysis of selection pressures

Normal B cells are selected based on their immunoglobulin's affinity for antigen, which can be visualised through the analysis of somatic mutation patterns in the *IGHV* (replacement versus silent). An increased frequency of replacement mutations in the

CDRs and a reduced frequency of replacement mutations in the framework regions which provide the structural backbone to the receptor, is indicative of antigen-driven selection. To determine whether our NGS derived sequences showed signs of antigen-driven selection, I used the online program BASELINe version 1.3 (Yaari et al., 2012). BASELINe is a statistical computational framework that quantifies selection from large-scale Ig sequence data by analysing the distribution of silent and replacement mutations in the CDRs and FWRs and comparing it to an expected frequency (Yaari et al., 2012, Uduman et al., 2011). The V gene sequences derived from the experimental data were inputted into the program in a single FASTA formatted file. This included the germline sequence which was placed above the sequences and contained an additional '>' in its title as an identifier to the program. The analysis output included identification of point mutations for each sequence, describing whether they were silent or replacement, and indicating their location in the variable region (CDRs or FWRs). This is compared against the expected frequency. *P*-values were also generated for both CDRs and FRWs indicating the selection strength which could be either positive or negative. The results from the multiple sequences were combined and gave a graphical output of the BASELINe analysis. An example is given below (Figure 2.3). The narrower the curve, the more confident the selection.

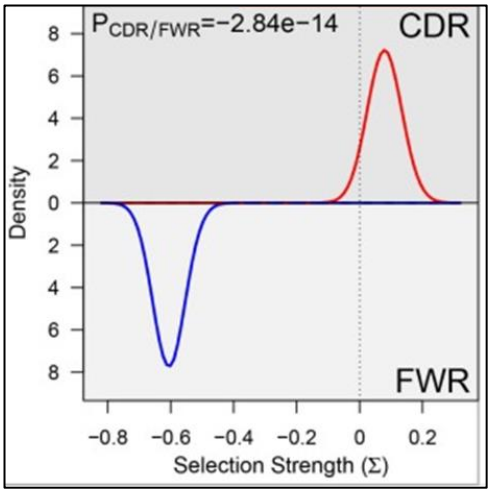


Figure 2.3: *Graphical output from BASELINe.* A value on the left-hand side of 0 (indicated by a dotted line) indicates a negative selection and a value on the right-hand side indicates a positive selection. This selection pressure profile graph is characteristic for antigen driven clonal expansion, with positive selection in the CDRs and negative selection in the FWRs. A negative $P_{CDR/FWR}$ indicates a more positive selection in CDRs compared to FWRs in clonally related sequences.

2.15 BCL2-IGH breakpoint analysis

As the t14;18 translocation occurs during the precursor B cell stage of development and is retained in tumour cells, the t14;18 breakpoint, which is unique to the clone, is often used to determine clonality between serial samples. This was to ensure that all successive events had arisen from a shared B cell precursor and were not *de novo* events. A qualitative analysis of the breakpoint was investigated through PCR and Sanger sequencing. The majority of breakpoints occur in the major breakpoint region (MBR), therefore a primer targeting a region of the *BCL2* gene which is upstream of this site was employed (Gribben et al., 1993). A consensus JH primer targeting the six germline JH segments was also utilised. The 5' to 3' sequence of primers is given below (Table 2.1). The PCR was performed using the HotStarTaq Master mix (Qiagen) with the PCR reaction detailed below (Table 2.2 and Table 2.3). A NTC was used in addition to a positive control (SU-DHL6) to evaluate sensitivity. Amplified PCR products were visualised on a 2% agarose gel, followed by clean-up and then sequenced using MBR and JH primers. The size of DNA bands differed between patients depending on the breakpoint location.

Primer	**5'-3' sequence**
MBR	CTATGGTGGTTTGACCTTTAGAGAG
JH	ACCTGAGGAGACGGTGACC

Table 2.1: *Primers used to amplify the BCL2-IGH gene rearrangements in temporal FL samples*. Primer sequences flank the major breakpoint region (MBR) of *BCL2* and JH region in the immunoglobulin heavy chain gene

Component	Volume (µl)
HotStarTaq Master Mix	5
Coral Red Dye	1
Forward Primer (10mM)-MBR	1
Reverse Primer (10mM)-JH	1
Nuclease-free water	1
DNA (50ng/µl)	2
Final Volume	*11*

Table 2.2: *BCL2-IGH breakpoint analyses PCR master mix*

Step	Temperature (^{o}C)	Time	
Initial denaturation	95	5 minutes	
Denaturation	95	30 seconds	**40 cycles**
Annealing	63	30 seconds	
Extension	72	30 seconds	
Final extension	72	7 minutes	
Hold	4	-	

Table 2.3: *PCR cycling conditions used to amplify the BCL2-IGH gene rearrangements*

The sequences were entered into BLAST (https://blast.ncbi.nlm.nih.gov/), a web alignment tool that finds regions of similarities between query nucleotide sequences and genomic databases. The top two hits included *BCL2* (Accession No. NC_000018.10) and *IGH* (Accession No. NC_000014.9). The fusion site was identified as the nucleotides at the interface of the BCL2 and IGH sequences which was unique to each patient.

2.16 Cloning

To analyse tumour subclones with differing *IGHV* sequences, *IGHV* PCR products were cloned using the TOPO TA cloning kit (Thermofisher). As the PCR products generated

using the proofreading enzyme are blunt-ended, 3' adenine overhangs were added for insertion into the plasmid vector. This was necessary as blunt-end cloning is ~100X less efficient than overhang cloning (Yao et al., 2016), resulting in fewer colonies in addition to the PCR product being inserted into the vector in the incorrect orientation. Following amplification of the *IGHV*, the PCR product was purified to ensure removal of proofreading DNA polymerase. A 10µl reaction mixture was set up in a standard PCR tube (Table 2.4).

Component	Volume (µl)
Purified blunt-ended PCR product (100ng)	2
dATP (1mM)	1
GoTaq Flexi DNA polymerase (5u/µl)	0.2
5 X GoTaq Reaction buffer (colourless)	2
Nuclease-free H_2O	4.8

Table 2.4: *Reaction mixture for adding 3' adenine overhangs to blunt-ended PCR product*.

The reaction mixture was incubated at 72°C for 30 minutes in a thermal cycler to add 3 A' overhangs to PCR products. The reaction mixture was placed on ice before directly proceeding to cloning.

TOPO cloning utilises the *Vaccinia virus* topoisomerase I which functions as both a restriction enzyme and a ligase. Topoisomerase recognises the sequence 5´-(C/T)CCTT-3´ which are found in the TOPO vectors and cuts at these sites, generating 3' overhangs of thymidine residues which then allows a single PCR amplicon with 3' adenine overhangs to ligate with the vector using ligation activity of topoisomerase I. These vectors are then transformed into E-coli which then replicate the recombinant DNA along with the host DNA, giving rise to bacterial colonies containing multiple copies of the DNA sequence, established from a single PCR amplicon. Different colonies contain different PCR amplicons which can then undergo sequencing to determine heterogeneity within the initial PCR product.

2µl of the reaction mixture described above containing the 3 A' overhang PCR product was added to 1µl salt solution (200 mM NaCl and 10mM MgCl2), 1µl of pCR®2.1-TOPO® vector (Figure 2.4) and 2µl of water. The cloning reaction was incubated for 5 minutes at room temperature to enable insertion of the PCR product into the vector. 2µl of the cloning reaction mix was added to one vial of One Shot® Chemically Competent E. coli and mixed gently before incubation on ice for 30 minutes. Cells were pre-treated with calcium chloride which enables close contact between the plasmid and the cells and facilitates plasmid uptake. The cells were then 'heat shocked' for 30 seconds at 42°C to enable plasmid entry by creating pores in the cell membrane. Immediate transfer of the tubes to ice results in restoration of the cell membrane potential. 250µl of room temperature S.O.C medium was added and tubes were shaken for 1 hour (200rpm) at 37°C to enable the recovery of E coli cells following the transformation process. 50µl and 100µl of medium were spread on two separate, pre-warmed LB agar plates containing ampicillin (50µg/ml) and X-gal (40mg/ml). Two plates at different volumes ensured that at least one plate would have well-spaced colonies which would be easy to pick. The pCR®2.1-TOPO® vector contains an ampicillin resistance gene (Figure 2.4) and therefore only E.coli which have been successfully transformed with the vector will survive on the ampicillin-coated plate.

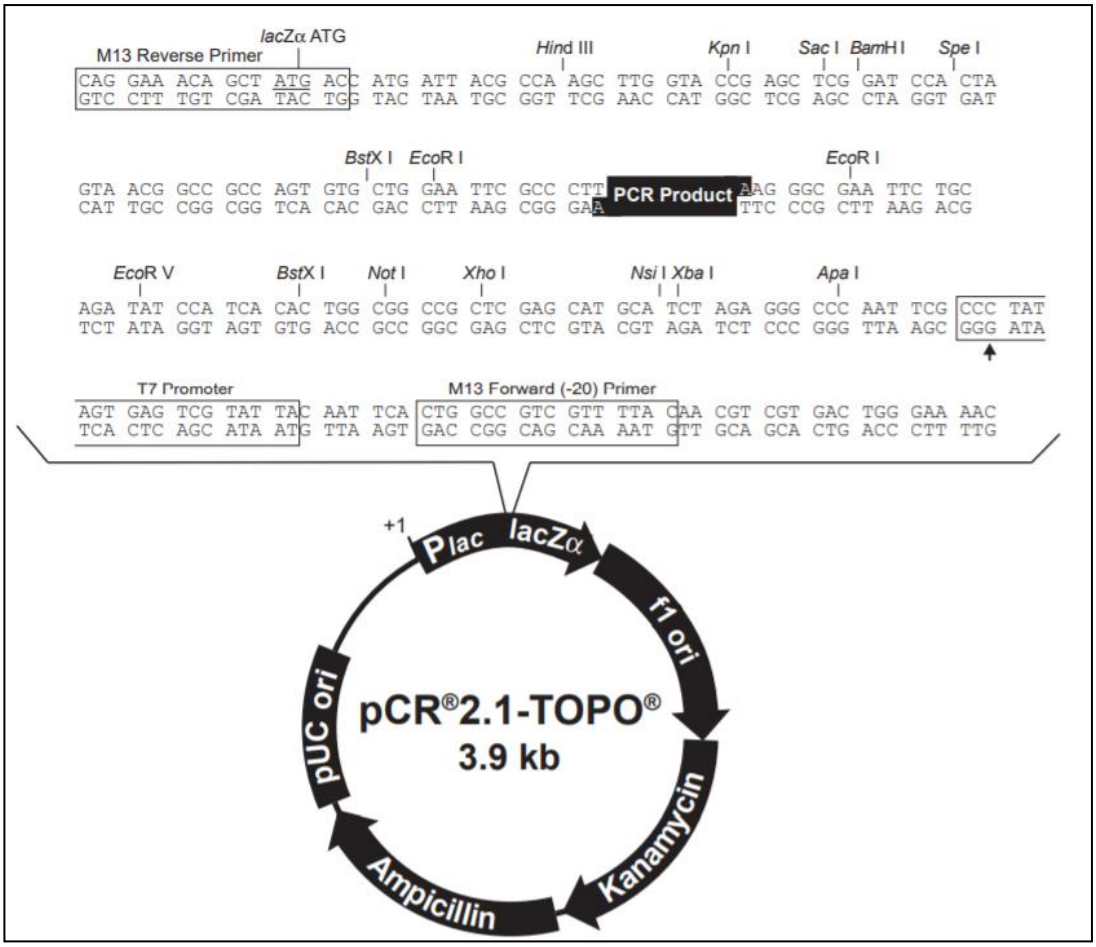


Figure 2.4: *pCR™2.1 TOPO® TA vector*. The vector has several restriction enzyme sites within the MCS including sequences that are recognised by topoisomerase I, which creates thymidine overhangs that are complementary to the adenine overhangs of the PCR product (highlighted), enabling successful incorporation of the DNA into the vector. The vector has ampicillin and kanamycin resistant genes and either of these antibiotics can be used in selection of E.coli transformed cells.

The vector also contains a LacZα gene that encodes for β-galactosidase, an enzyme that metabolises lactose. As the PCR product is inserted into a multiple cloning site (MCS) within the LacZα gene, the production of a functional β-galactosidase is inhibited (Figure 2.4). As X-gal is a substrate for β-galactosidase, and forms an insoluble blue pigment upon its hydrolysis (Figure 2.5), we were able to differentiate which vectors have successfully taken up the PCR product by the colours of the colonies. Blue colonies have no insert as the LacZα gene had not been disrupted whereas white colonies indicated successful incorporation of the insert as X-gal has not been hydrolysed by β-galactosidase (Figure 2.5).

Individual white colonies representing different PCR amplicons were picked. The vector also contains M13 sequences within the multiple cloning site (Figure 2.4), enabling direct sequencing of individual colonies using either using M13 forward (CATTTTGCTGCCGGTC) or M13 reverse (GTCCTTTGTCGATACTG) primers. PCR amplification of the colonies was carried out to ensure enough sample for sequencing. Each colony was individually resuspended in 48µl of Platinum™ SuperFi™ PCR Master Mix (Thermofisher) and 1µl of both forward and reverse M13 primer. The reaction was incubated for ten minutes at 94°C before amplification for 35 cycles of denaturation, reannealing and extension at 95°C, 56°C and 72°C, respectively. This was followed by a final extension of seven minutes at 72°C before a 4°C hold. 1x gel loading dye was added to the PCR product before sample underwent agarose gel electrophoresis. Bands were visualised under UV light and purified before sending for sequencing (Germany).

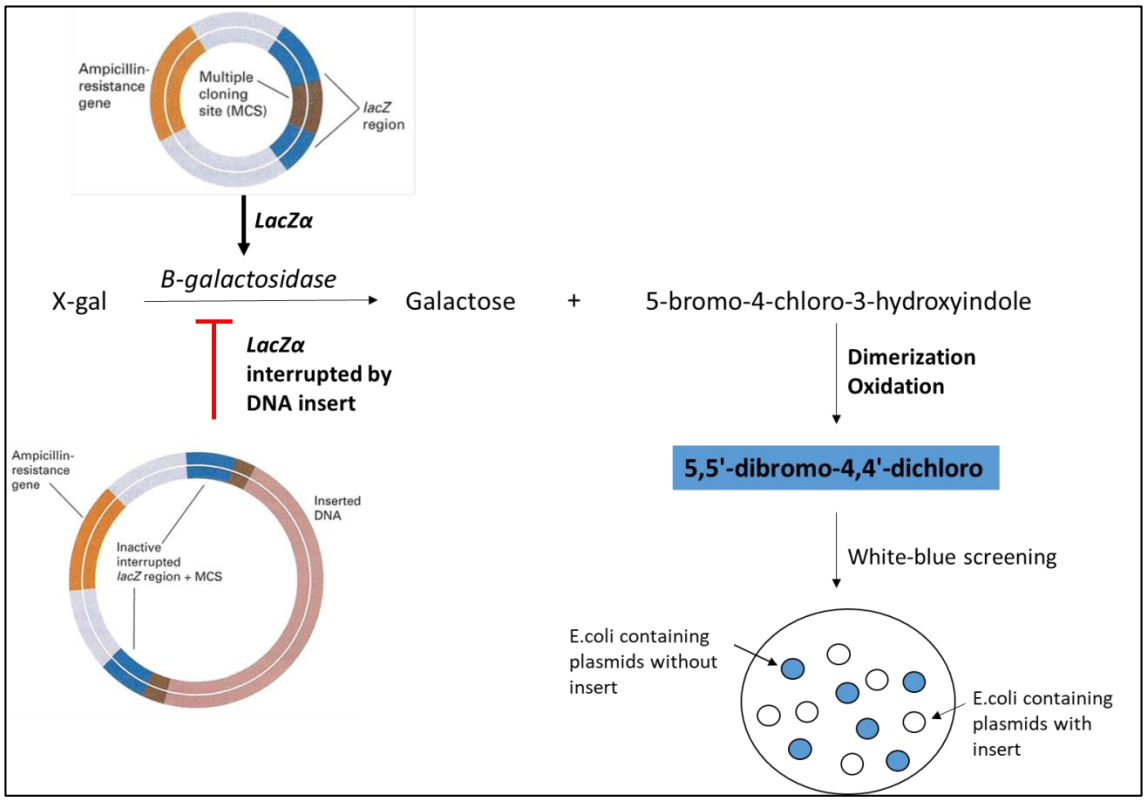


Figure 2.5: *Visualisation of successful incorporation of the DNA in the plasmid vector.* X-gal is hydrolysed by β-galactosidase, an enzyme encoded by the LacZα gene, to produce galactose and 5-bromo-4-chloro-3-hydroxyindole. 5-bromo-4-chloro-3-hydroxyindole undergoes oxidation and dimerisation to produce the insoluble blue pigmented precipitate, 5,5'-dibromo-4,4'-dichloro. When the LacZα gene is disrupted by DNA insert, β-glactosidase is not produced and therefore cannot catalyse hydrolysation of X-gal. Determing which plasmids have taken up the DNA insert and which haven't can be visualised by blue-white screening of colonies on an agar plate which contains X-gal.

2.17 Targeted amplicon sequencing

Amplicon sequencing enables ultra-deep sequencing of PCR products to enable variant identification at a much higher sequence depth and lower cost than WGS. This is advantageous in the detection of clonal and importantly, subclonal variants.

Following PCR amplification of the *IGHV*, 100ng of purified DNA was sent to Genewiz (UK) for paired end 2x250bp sequencing using the Illumina Miseq platform. Additional quality control (QC) was performed using the D1000 ScreenTape assay to determine suitability of the DNA for the NGS platform. This required 1-2 µl of DNA sample to undergo automated electrophoresis and gel imaging to assess the quantity, size and integrity of DNA using TapeStation analysis software. This step verified the insert size was as expected (250-350bp) and there was no contamination by larger or smaller DNA fragments.

2.17.1 Preparation of libraries

Samples underwent multiplexed sequencing in which DNA libraries are tagged with unique indexes made of eight nucleotides enabling multiple samples to be pooled onto a single flow cell and sequenced simultaneously. This offers the benefit of increased experimental throughput whilst maintaining low error rates due to tag identification. Indexes are added to DNA fragments during adapter ligation, carried out using TruSeq DNA Library Preparation Kit (version 2). Firstly, the 3' ends of the DNA fragments were adenylated (addition of a single adenine nucleotide), a process known as A-tailing, to prevent fragments from ligating to each other. The 5' ends of the DNA fragments are phosphorylated, which enables successful ligation. Adapters are partial double-stranded molecules containing the full complement of sequencing hybridization sites, including read 1 and read 2 specific sequencing primer sites, index primer site and complementary sequences to the flow cell (Figure 2.6). The 3' end of the adapter contains a thymidine nucleotide overhang which is complementary to the adenine overhang on the DNA fragment, resulting in ligation of the adapter to the DNA. PCR was used to selectively enrich DNA fragments with adapters ligated to both ends. The PCR primers anneal to

the ends of the adapters and amplify the amount of DNA in the library. Final indexed DNA libraries are normalised to 10nM and pooled in equal volumes in preparation for cluster generation.

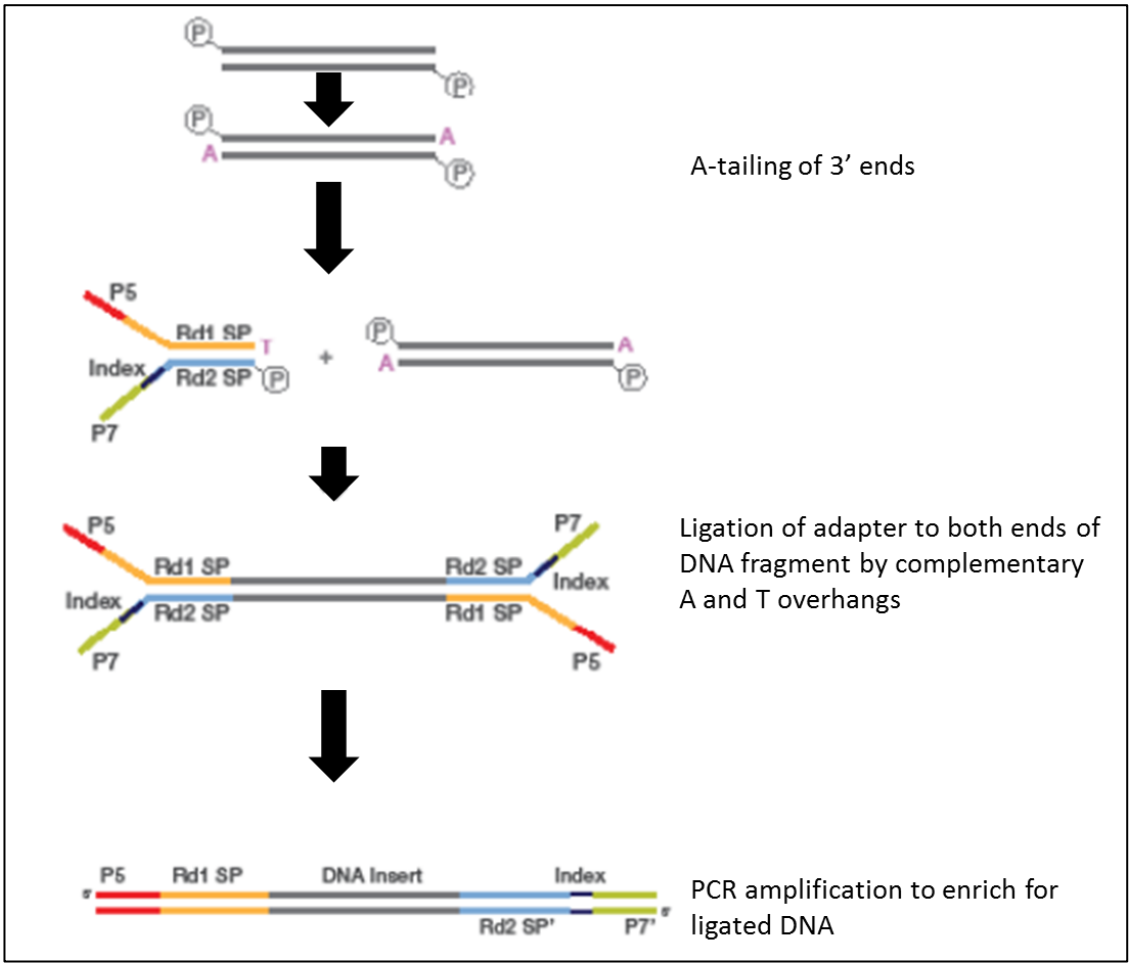


Figure 2.6: *DNA library preparation*. Single adenines were added to the 3' ends of amplicons. A corresponding single thymidine nucleotide on the 3' end of the adapter provides a complementary overhang for ligating the adapter to the fragment. The adapter consists of the Read 1 sequencing primer site (Rd1 SP), Read 2 sequencing primer site (Rd2 SP), an index sequencing primer site, and two sequences (P5 and P7) that are complementary to the two types of primer attached to the surface of the flow cell used for cluster generation through bridge amplification. Only DNA fragments which have adapters incorporated on both ends are able to form clusters on the flow cell.

2.17.2 Cluster generation and sequencing by synthesis

Single-stranded DNA libraries are applied to a flow cell coated with oligonucleotides that are complementary to those found on the adapter region of the DNA, specifically P5 and P7 oligonucleotides. When DNA binds to the complement oligonucleotides on the flow cell, cluster generation occurs in which approximately a million copies of the template strand are created through bridge amplification. Reverse strands are then washed away and the 3' end of the forward strands are blocked before they undergo sequencing. The sequencing primer binds to it complementary sequence near the 3'end followed by incorporation of fluorescently labelled ddNTPs by a polymerase. Only one fluorescently tagged ddNTP is added to the growing chain per cycle based on the sequence of the template. All identical strands in a cluster are sequenced simultaneously. A light source excites the fluorophore, resulting in an emittance of a fluorescent signal in the cluster after each round. The fluorophore is then cleaved and the 3' hydroxyl group is regenerated with a reducing agent, enabling the next ddNTP to be added and the process is repeated. This is known as sequencing by synthesis (SBS). After the Read 1 product is produced, it is washed away. The index primer is then introduced and anneals to the same strand to produce the 8-bp index read which is used to identify samples during downstream analysis (Figure 2.7). To produce a paired-end read, the 3' end of the strand is unblocked and the template folds over and bind the second oligonucleotide on the flow cell surface, resulting in the regeneration of the complementary (reverse) strand (Figure 2.7). Following linearization of the DNA, the forward strands are removed and the complementary strands are sequenced following binding of the read 2 sequencing primer to its site on the strand to generate the paired read 2 product. Sequences were de-multiplexed according to their index read sequences to identify sample origin.

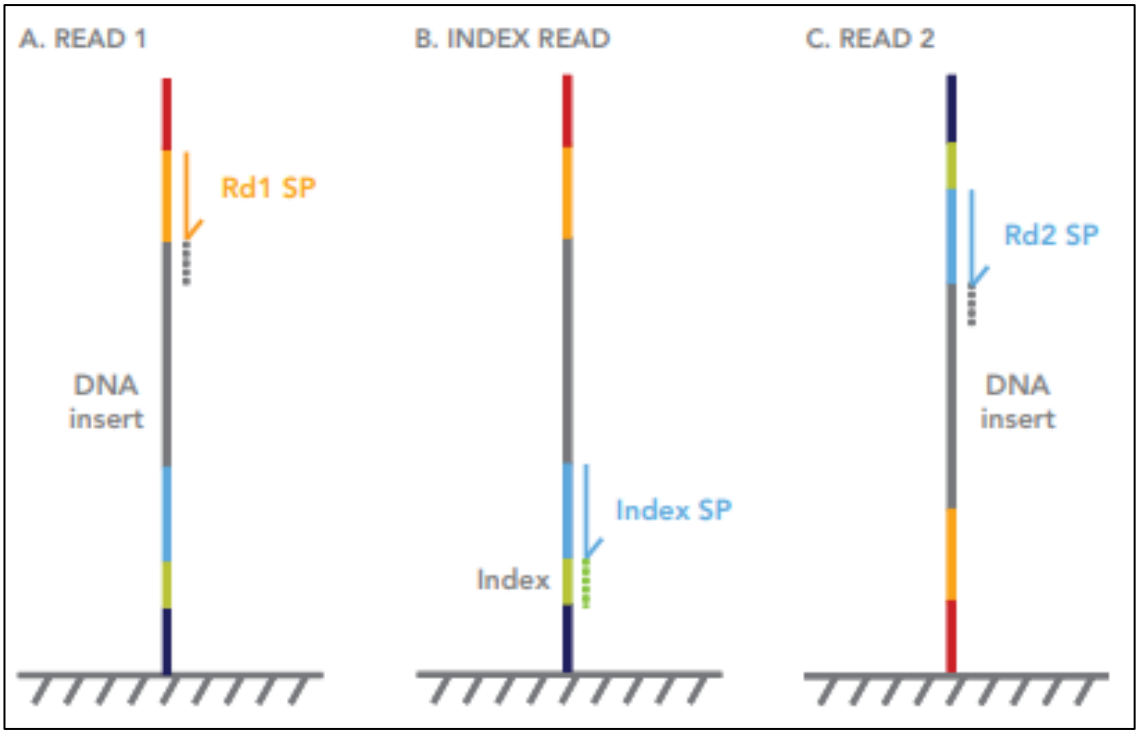


Figure 2.7: *Multiplexed paired-end sequencing process*. (A) Forward strands attached to the flow cell are used to create Read 1 products. Complementary primers to the Rd1 SP sequence bind and polymerases add fluorescently labelled ddNTPs through SBS. (B) The Read 1 product is then removed and an index primer then binds to the index sequence to produce an 8bp index read. This read is then removed, followed by the forward strand forming a bridge with the second set of oligonucleotides on the flow cell (not shown) and producing a complementary (reverse) strand (C). The forward strands are removed, leaving the reverse strand to undergo sequencing following primer binding to the Rd2 SP site.

2.17.3 Phasing

An issue of the Illumina is that as the number of cycles increases, the quality of base calls decrease, resulting in the end of reads being less accurate than the beginning of reads. This is because the SBS process relies on sequencing millions of identical molecules (a cluster) simultaneously, with each cycle adding one nucleotide at a time. However, a small proportion of molecules in the cluster will not undergo extension and remain on the base of the previous cycle, becoming out-of-sync to the rest of the cluster. These molecules will pollute the light signal emitted from the fluorophores coupled to the newly incorporated nucleotide in the majority of molecules. This is known as phasing (Figure 2.8). The proportion of molecules affected by phasing increases with cycle

number, hindering the correct identification of bases in longer reads (Erlich et al., 2008, Kircher et al., 2009). To reduce base inaccuracy in the CDR3 region of the *IGHV*, paired-end sequencing was employed, in which the forward and reverse strands were sequenced (Figure 2.9). The reads generated were 250bp in length. Based on a 300bp *IGHV* fragment, the reads overlapped by 200bp, a sufficient amount to enable merging of reads, yielding sequences of higher quality that accounted for sequencing errors.

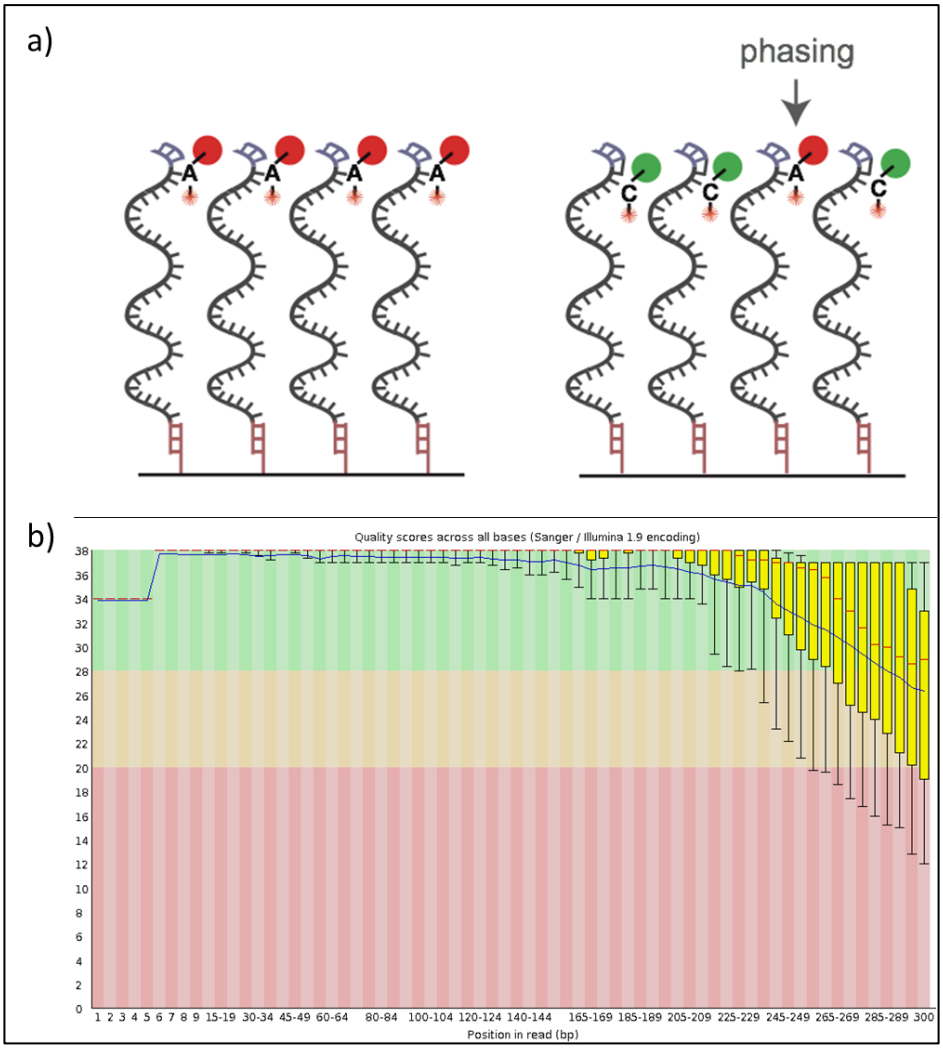


Figure 2.8: *Reduced base quality during ongoing cycling using Illumina technology*. a) During one cycle, clusters incorporate a single fluorescently labelled nucleotide (left panel). When molecules of a cluster do not incorporate a new nucleotide during the next cycle (right panel), they are one base behind the rest of the cluster (e.g. out of phase), resulting in pollution of the light signal used to generate base calls. b) Typical range of quality scores (y-axis) across all bases at each position in the read (x-axis). Yellow boxes represent the inter-quartile range (25-75%), the red line represents the median value, and the blue line represents the mean quality.

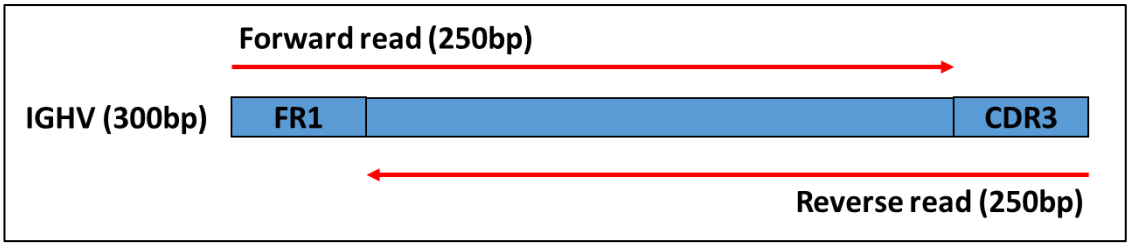


Figure 2.9: *Paired-end read length in relation to target DNA fragment length*. The reads overlap by 200bp based on a 300bp fragment.

2.17.4 PhiX Spike-in

Because the B cell clones carry a highly similar sequence and therefore low base diversity, a 30% PhiX Spike-in was included to allow accurate base calling and ensure the complexity of the run. The PhiX genome library has a diverse base composition with all four nucleotides present in similar quantities (45% GC, 55% AT), compensating for the low base diversity in the samples, and therefore assisting with template registration on flow cells and as a calibration control for phasing and pre-phasing calculations in addition to registering the overall performance of the sequencing run.

2.18 NGS data processing

Data processing was carried out by Genewiz including de-multiplexing, adaptor trimming and removal of low quality bases located at read ends. Paired reads were merged using the BBMerge script with the default setting requiring a minimum insert of 35bp and a minimum identical overlap of 12bp to allow merging. Reads that were not merged were discarded. Unique sequences (excluding primer sequences) were identified and their abundance was quantified as a count number.

2.19 NGS processing: IMGT/H*IGHV*-Quest

Sequences with counts of ≥10 were selected for analysis using IMGT/H*IGHV*-Quest (Alamyar et al., 2012), the high-throughput version of IMGT which is capable of analysing up to 500,000 immunoglobulin sequences in a single run. Sequences were first converted to the FASTA format and entered into the search page as a text file. Output files included the Summary file which contained information for each sequence in relation to its functionality, V, D and J genes and alleles with closest homology, junction analysis and FR-IMGT and CDR-IMGT lengths, all according to IMGT ontology and classification. Because healthy B cells also display immunoglobulins of the *IgH-V3* family, there was a heterogeneous level of contaminating sequences from normal B cells in our sequencing data. However, as we previously identified the major clone for each sample through heteroduplex analysis, tumour related sequences were identified as having identical *V*, *D*, *J* genes and CDR3 length to the major clone. Following their selection, tumour related reads underwent further filtering regarding functionality and CDR lengths (Figure 2.10).

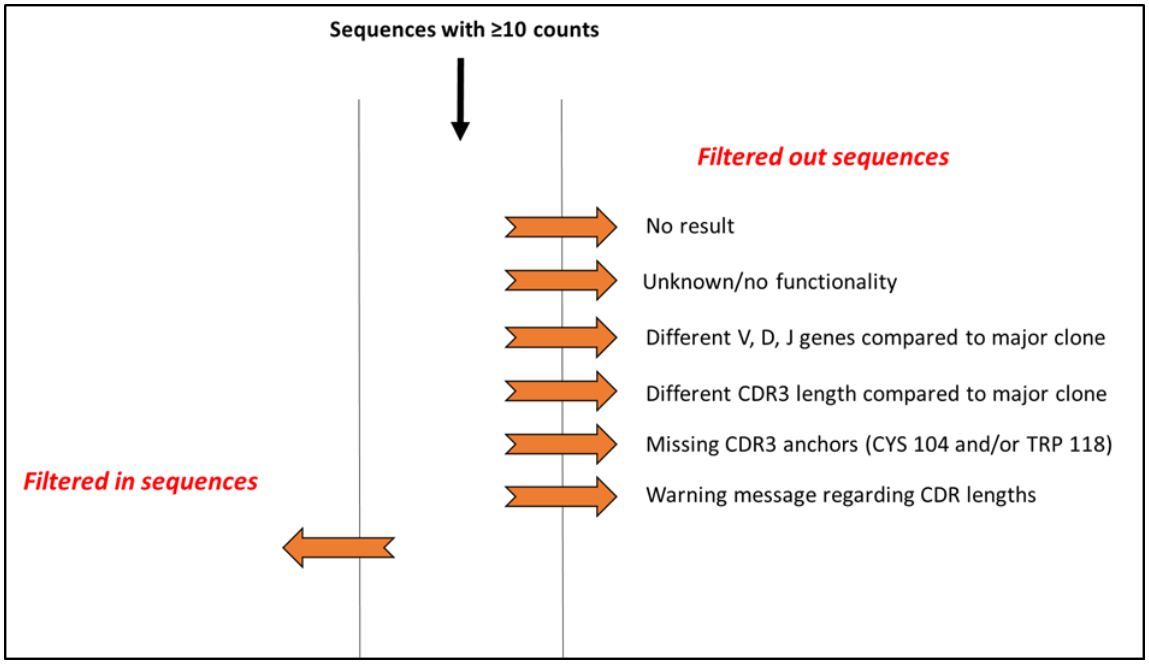


Figure 2.10: *Parameters used in selection and filtering of unique sequences*.

2.22 NGS processing: Editing

To ensure sequences covered an identical region, multiple sequence alignment program (MAFFT) was used (usegalaxy.org) to compare regions of similarity. Sequences with deletions or insertions were omitted due to our focus being on single nucleotide polymorphisms (SNPs) related to the SHM process. Sequences were trimmed to the length represented by the majority of sequences to retain diversity. Duplicate sequences were found and omitted, resulting in unique sequences differing from each other by ≥1 nucleotide. This provided us with our final clonal repertoire of *IGHV* sequences which were used for lineage tree construction below.

2.20 *IGHV* NGS data for serial FL samples

To validate findings from our *IGHV* sequencing platform, raw *IGHV* sequencing data files deposited in databases were utilised and underwent processing.

Tumour-related reads covering the *IGHV* gene for three patients over different disease events were downloaded from the European Nucleotide Archive (ENA) using the accession numbers PRJEB9334 and PRJNA240336 (Jiang et al., 2014, Carlotti et al., 2015). Reads were generated using the Roche 454 Life Sciences Genome Sequencer FLX or Illumina Miseq platform. Temporal biopsies were taken either from the same lymph node or distinct lymph nodes, which will be detailed in the relevant results chapter.

2.21 *IGHV* sequences for composite HL and NHL lymphomas

IGHV sequences for Reed Sternberg and FL cells in composite Hodgkin's lymphoma (HL) and FL disease from three patients were downloaded from the ENA using the following accession numbers; AJ409171, AJ409173-AJ409186 (Kuppers et al., 2001), AJ011134 and AJ011135 (Brauninger et al., 1999). HL and FL cells were derived from the same biopsy site, which will be detailed in the section 3.4.17.

2.23 Statistical Analysis

Statistical analyses were performed using GraphPad Prism version 5.4 (GraphPad Software Inc., La Jolla, CA, USA). Two way ANOVA was used to determine if the differences between the number of N-gly positive subclones and N-gly negative subclones within each disease event was greater than you would expect to see by chance. Statistic tests were considered significant at $P \leq 0.05$ (*). A *p*-value of ≤0.01 was identified by two asterisks (**) and a *p*-value of ≤0.001 was identified by three asterisks (***).

2.24 Lineage trees

To visualise clonal expansion and the relationship between individual subclones, lineage trees were constructed using the algorithm of the IgTree program (Barak et al., 2008). IgTree is specifically designed for Ig sequences, calculating the distance between each pair of input sequences to find possible ancestor–progeny relationships, with the putative germline sequence being the root of the tree.

To ensure lineage trees accurate represented the evolutionary hierarchy of lymphoma-derived clones, the *V*, *D* and *J* gene sequences need to be included as SHM can occur across the entire variable region. However, the CDR3 region poses a challenge; during somatic recombination, P- and N-nucleotides are added at the junctions between the gene segments and are therefore not encoded by genes. The lack of an experimentally derived germline *IGHV* sequence means the original P- and N- nucleotides inserted during somatic recombination cannot be identified and therefore within this region we cannot determine which nucleotides are present because of SHM or somatic recombination, resulting in inaccurate lineage trees. Therefore, P- and N- nucleotide regions identified by IMGT/V-QUEST in experimentally derived sequences were changed to 'N' whilst nucleotides assigned to V, D or J genes were retained. The IgTree program ignores these 'N' regions, generating lineage trees based only on mutational profiles of V, D and J derived sequences. N was also inserted into corresponding sites in the putative germline sequence.

Following editing and alignment of experimental *IGHV* sequences and the putative germline sequence using MAFFT (usegalaxy.org), the output FASTA file was converted to a .PIR file. This file type was used as the input file for the IgTree program.

The output tree file created is a .vsdot file which is visualised by Graphviz 2.38 (Figure 2.11). Every distinct sequence is set as a node within the tree, separated from its ancestor by one mutation, unless otherwise stated on the connecting branch (Figure 2.11). If a node has no descendants, it is set as a leaf. White nodes represent subclones which are artificially added by the IgTree program to connect experimental subclones. The program traces the minimal steps separating experimental sequences from each other and the putative germline, with the assumption that this is the naturally occurring scenario. It can also account for reversion mutations (point mutation that restores the original nucleotide). The program uses a heuristic algorithm to reduce the running time of the program and therefore short sequences (~300bp) and closely related sequences are best suited to this program. One advantage is that trees are non-binary, reflecting Ig evolution in which several B clones can arise from one cell.

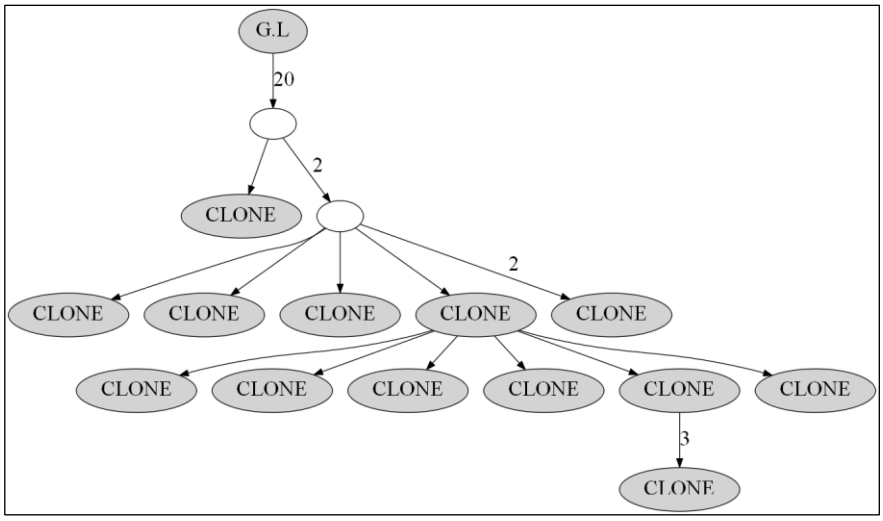


Figure 2.11: *Example of a lineage tree created using IgTree and Graphviz.* The root of the tree marks the known Ig germline (G.L) sequence to which experimental Ig sequences are compared. Each node represents a unique sequence and therefore a subclone. Clear nodes represent clones that have not been found from the sequencing data but are implied to exist. Nodes are separated by at least one mutation, unless indicated on the branches. Internal nodes give rise to other subclones whereas nodes, which do not daughter clones, are set as leaves. The tree enables the visualisation of evolution of mutations/subclones within the Ig variable region, making it a useful tool for identifying tumour populations and progression.

2.25 Immunohistochemistry

Immunohistochemistry (IHC) is able to determine the tissue distribution of a specific antigen *in situ* based on a chromogenic signal. Briefly, the indirect method of IHC utilises a primary and secondary antibody to amplify detection of an antigen. The primary antibody binds to the target antigen on the tissue via the Fab regions, forming the first layer. The secondary antibody which is labelled with an enzyme binds to the primary antibody. Following addition of the enzyme substrate, a coloured precipitate is deposited at the site of antigen (Figure 2.12).

As DC-SIGN and MR are expressed on macrophages and dendritic cells, we included markers to identify these cell types. For macrophages, the pan-marker CD68 and M2 macrophage marker CD163 were used. CD21 was used to identify follicular dendritic cells to determine whether expression of lectins occurred in this germinal centre-based dendritic cell type. CD21 is also a marker for mature B cells, however as these lectins are not expressed on B cells, CD21+DC-SIGN+/MR+ staining was considered to be restricted to follicular dendritic cells.

3µm thick whole sections were cut from tissue blocks and mounted onto glass slides. Slides were incubated in a 60^{o}C oven overnight to ensure effective adhesion of the tissue to the surface and to soften the wax. Slides were prepared for immunostaining by complete removal of the paraffin using xylene. Slides were immersed in two rounds of xylene for five minutes each. Xylene was then removed and the section re-hydrated through graded washes with industrial methylated spirit (IMS) and water. Endogenous peroxidase activity was blocked by incubating slides in 2% H_2O_2 in IMS solution for two rounds of five minutes. This was to minimise non-specific background staining as horseradish peroxidase-conjugated secondary antibodies were later used as the substrate for colour precipitation and positive staining. Formaldehyde fixation results in cross-linking between proteins, resulting in a change to the tertiary structure and potential hindrance of epitope recognition by the primary antibody. To remove methylene bridges between proteins, heat-induced epitope retrieval (HIER) was used. HIER was performed for ten minutes in a pressure cooker (400^{o}C) containing three litres

of citrate based antigen unmasking solution (Vector Laboratories H3300) at a 1:100 dilution in water. Slides were then rinsed in running water for five minutes to remove residual unmasking solution and placed in wash buffer (DAKO), a tris-buffered saline with tween (TBS-T). From this point on, slides were kept wet to avoid non-specific binding of antibody. The area around the section was dried and drawn around using a hydrophobic pen. Wash buffer was quickly applied to the section to prevent drying out of the tissue. Optimal primary antibody dilutions (using Zytomed Antibody diluent) were determined by titration around the dilution recommended from the datasheet (Table 2.5). Primary antibody directed against the antigen of interest was applied to the section for a forty minute incubation.

Antibody (company, catalogue number, clone)	**Validated working dilution**
CD10 (Novocastra, NCL-CD10-270, mouse monoclonal IgM, clone 56C6)	1/100
CD68 (Dako, M0814, mouse monoclonal IgM, clone KP1)	1/8000
CD21 (Novocastra, NCL-CD21-2G9, mouse monoclonal IgM, clone 2G9)	1/100
CD163 (Novacastra, NCL-L-CD163, mouse monoclonal, clone 10D6)	1/250
DC-SIGN (LSBio, LS-B479, Rabbit IgG polyclonal)	1/3500
MMR (R&D, MAB25341, mouse monoclonal IgG, clone 685645)	1/200

Table 2.5: *Primary antibodies and dilution factors.*

To visualise antibody-antigen interactions a peroxidase-labelled VIP system (Super Sensitive Polymer-Horseradish Peroxidase IHC Detection System; BioGenex) was used. A secondary antibody, directed against the species in which the primary antibody was raised in, was conjugated to horseradish peroxidase (HRP)-polymers. Following a 30 minute incubation with the secondary antibody, hydrogen peroxide and the chromogen VIP (Vector Laboratories) are incubated with the section which react with the HRP polymer to form a purple precipitate. Tissue sections were counterstained in Mayer Haematoxylin (Gill's II) which binds to lysine residues on nuclear histones, resulting in a blue stain of the nuclei. Slides were then dehydrated in IMS and xylene before mounting with DPX xylene based mountant and left to dry overnight in the fume cupboard. Slides

were scanned using the Pannoramic 250 High Throughput scanner to produce high quality brightfield images. Quality of scanned slides were checked on the Pannoramic Viewer program (Version 1.15.4).

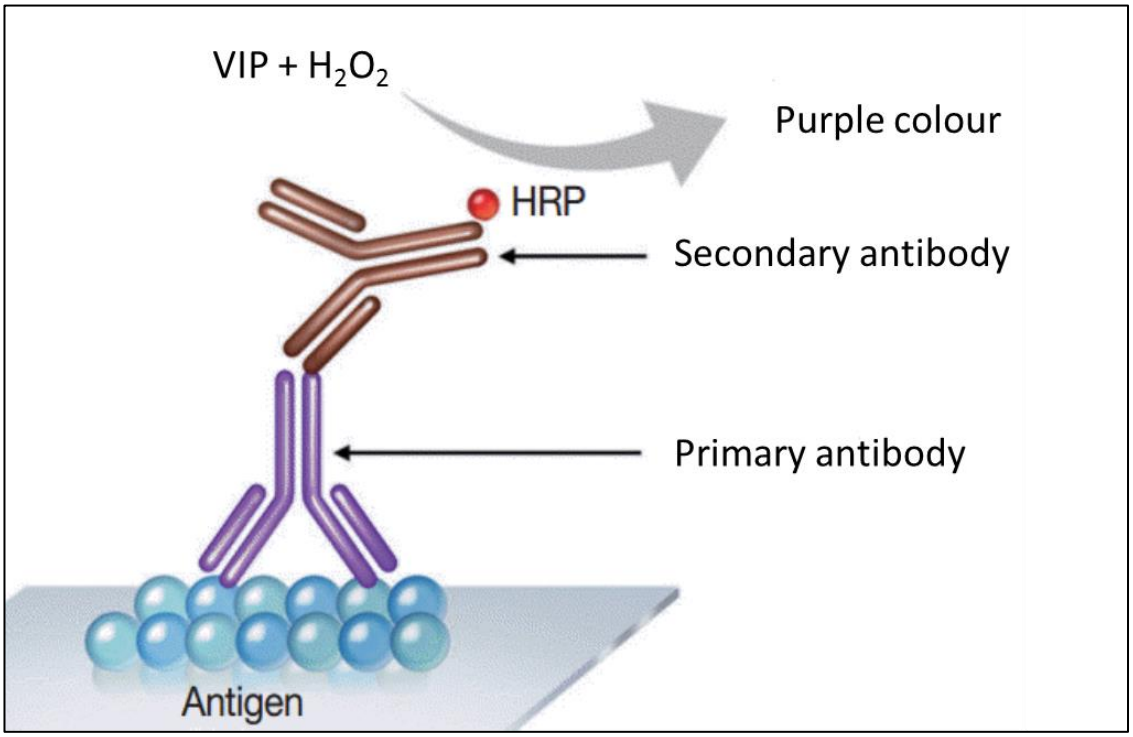


Figure 2.12: *Indirect IHC*. A primary antibody which recognises epitopes on the antigen of interest binds to the surface via its Fab domains. A secondary antibody (conjugated to HRP) which has been raised against the host species of the primary antibody then binds to the unlabelled primary antibody. The addition of VIP and hydrogen peroxide act as substrates for the HRP enzyme, resulting in precipitation of an insoluble purple colour deposit at the site of the antigen. Although not shown here, two or more secondary antibodies can bind to a single primary antibody, resulting in signal amplification and detection of weakly expressed antigens.

2.26 Principles of multiplex IHC

Analysing the distribution of specific antigens within a tissue enables an important insight into the tumour microenvironment and biology of disease. Historically, this was performed on serial tissue sections, each stained with a different marker. Yet due to the three dimensional structure of the tissue, serial sections do not capture an identical array of cells leading to a lack of information regarding cellular co-expression and spatial

relationships between cells. Multiplex IHC offers simultaneous visualisation of multiple antigens on the same tissue section resulting in a more comprehensive characterisation of the microenvironment. This can be performed simultaneously, in which sections are stained with a variety of antibody-chromogen pairs that are visualised by either HRP or alkaline phosphatase (AP) (Table 2.6). To avoid secondary antibody cross-reactivity, primary antibodies must be from different host species (e.g. rabbit, mouse). Furthermore, while this method can be useful for different cell types, the co-localisation of antigens may be difficult to interpret by overlapping chromogens at a specific site. This limits multiplexing capacity and usually no more than three antigens are visualised this way. These issues can be overcome by sequential multiplex staining, in which the stain is stripped from the section and then re-probed with a different primary antibody. A technique developed and validated in our department by Andrew Clear (Ball et al., 2017) enables up to six repeated re-probes on whole tissue sections with very little amendment to the protocol described in section 2.25.

Chromogen	Catalytic agent	Deposition colour
DAB	HRP	Brown
DAB + Ni	HRP	Black
AEC	HRP	Red
VIP	HRP	Purple
NBT/BCIP	AP	Deep blue
Vulcan Fast Red	AP	Red
Vector Black	AP	Black
Nova Red	HRP	Deep red
TMB	HRP	Blue

Table 2.6: *Chromogens used in multiplex IHC (brightfield).* DAB (3,3'-Diaminobenzidine), Ni (Nickel), AEC (3-Amino-9-ethylcarbazole), NBT/BCIP(Nitroblue tetrazolium and 5-Bromo-4-chloro-3-indolyl phosphate), TMB (3,3',5,5'-Tetramethylbenzidine), HRP (Horseradish peroxidase), AP (Alkaline phosphatase). Table taken from Stack *et al*, 2014 (Stack et al., 2014).

2.27 Sequential multiplex IHC: stripping and re-probing

Following high-resolution scanning, slides were placed overnight in xylene, until coverslips could be easily removed without damaging the section. As the slides were dehydrated and mounted with a xylene based solution (DPX mountant, Sigma), rehydration was required before stripping, using graded IMS and water washes. Following a five minute incubation in running water, slides were placed in a pressure cooker for ten minutes containing 3L of citrate based antigen unmasking solution (Vector Laboratories H3300) at a 1:100 dilution in water for HIER. After cooling and rinsing in water, slides were checked by eye against a white background to detect any colour, indicative of residual staining. If no colour was detected, slides were kept in wash buffer and underwent the same protocol as above (Section 2.25) to detect the next antigen. This process was repeated until all six stains were scanned. To minimise tissue degradation with repeat rounds of stripping, boiling time was optimised to sevenminutes, which was found to be the shortest incubation time necessary to effectively strip the tissue.

To verify that primary antibodies could be stripped from the tissue, whole tonsil sections were used to test each antibody (Figure 2.13). Following stain and strip, the slide was analysed for any residual antibody by using the peroxidase-labelled VIP system (Super Sensitive Polymer-Horseradish Peroxidase IHC Detection System; BioGenex) (Figure 2.13). Residual staining indicated incomplete removal of the antibody. This is exampled for CD21. This helped decide the optimal sequence of re-probing, with CD21 being the last stain and therefore issues regarding effective stripping were addressed and minimised. To ensure that antigen stability was conserved through several rounds of repeated stain/strip cycles, two lymph node sections were subjected to the five rounds of cycles, in the following antibody order; DC-SIGN, CD10, CD163, MMR, CD68, CD21.

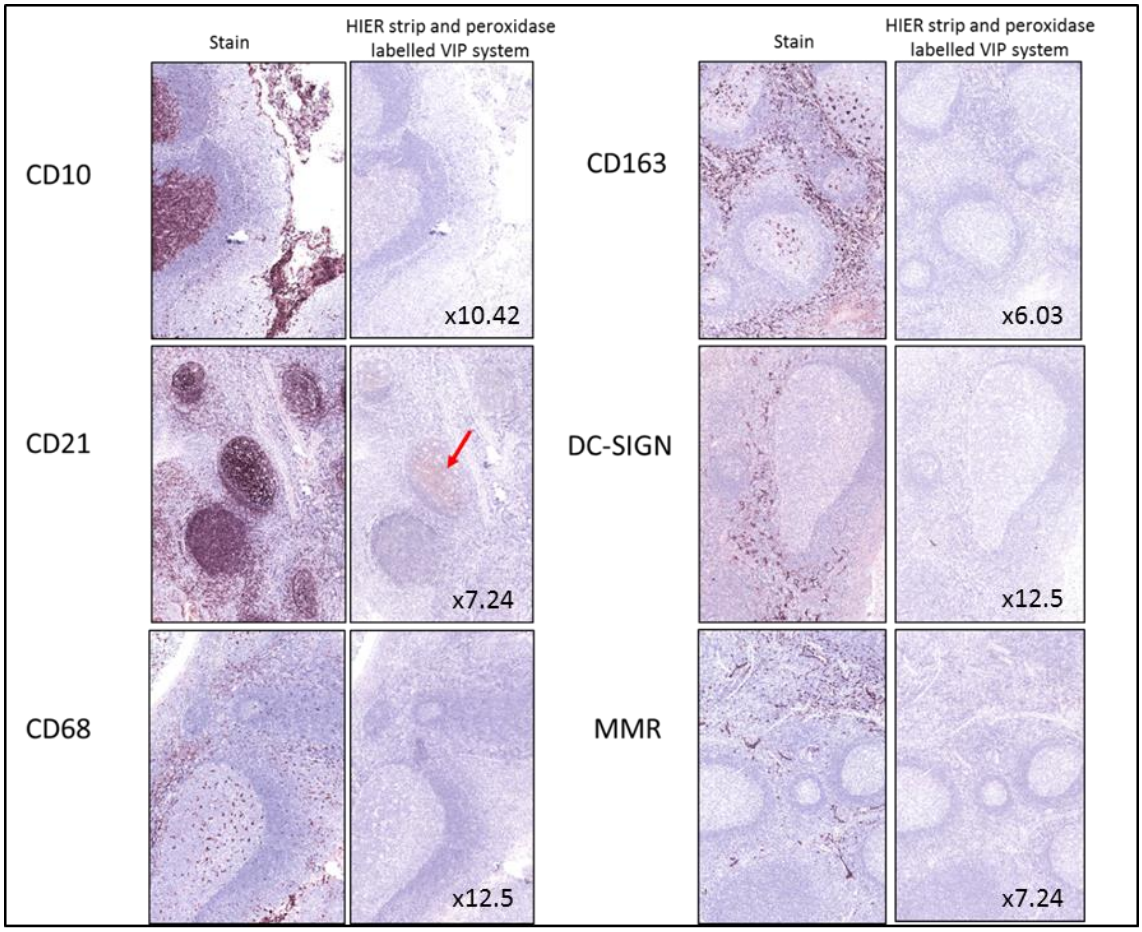


Figure 2.13: *Antibody test to determine the stripping efficacy*. Tonsil sections were stained with a primary antibody (CD10,CD21, CD68, CD163, DC-SIGN, and MMR) following the protocol in Section 2.25 (left columns of both panels). Following scanning, slides underwent HIER stripping (described above) and were exposed to a secondary antibody conjugated to HRP. Hydrogen peroxide and VIP were added to the slide and any deposit of chromogen on the tissue was indicative of residual primary antibody and ineffective stripping. This was seen for CD21, with residual chromogen indicated by the red arrow. CD21 was therefore decided to not be stripped from the tissue and therefore was the last antibody deposit in the stain/strip cycle. Magnification is indicated for each antibody image.

2.28 Digital image analysis

To determine the identification of DC-SIGN/MR-expressing cells required analysing the co-localisation of such biomarkers with cell specific biomarkers (CD68, CD21 and CD163). To achieve this, the Tissuealign™ analysis module in the digital image analysis software Visiopharm (Hoersholm, Denmark) was used to perform automated alignment of the six sequentially stained images. Correct alignment was critical to accurate

downstream quantitative analysis and therefore the alignment was checked by eye to ensure correct mapping of cells. Any misalignments were manually corrected.

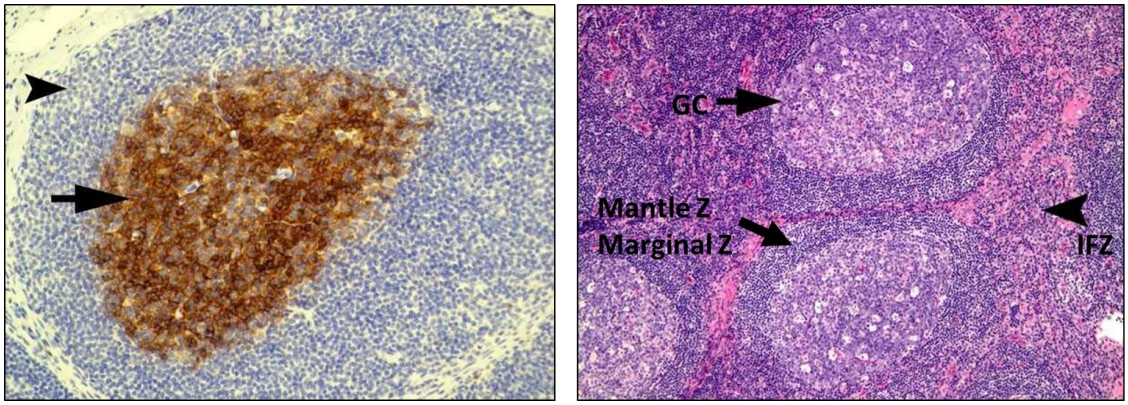


Figure 2.14: *Normal histology of lymph node*. Left panel) All germinal centre B cells express CD10 (brown stain indicated by arrow), with mantle zone B cells being CD10 negative (arrowhead). Right panel) H&E stain showing germinal centre (GC) surrounded by outer rings of mantle (Mantle Z) and marginal zone (Marginal Z) B cells. The interfollicular zone (IFZ) is indicated. Images adapted from www.pathpedia.com. Magnification and size unknown.

To identify the location of DC-SIGN/MR-expressing cells within the tumour microenvironment, an analysis protocol package (APP) was designed in Visiopharm to recognise and annotate GC and non-GC regions in all patient and control samples. Germinal centres are normally surrounded by a ring of small, dark, tightly packed B cells which comprise the mantle zone, followed by the marginal zone (Figure 2.14). In FL, neoplastic follicles can be back to back with the mantle zone usually attenuated or absent, making germinal centre identification difficult from the interfollicular zone. Therefore, CD10 staining was used to train the APP to identify GC regions as the biomarker is specifically expressed on both normal and neoplastic germinal centre B cells. Non-CD10 regions were allocated 'Tissue' and regions of no cells were designated 'Background' to exclude their analysis downstream (Figure 2.15).

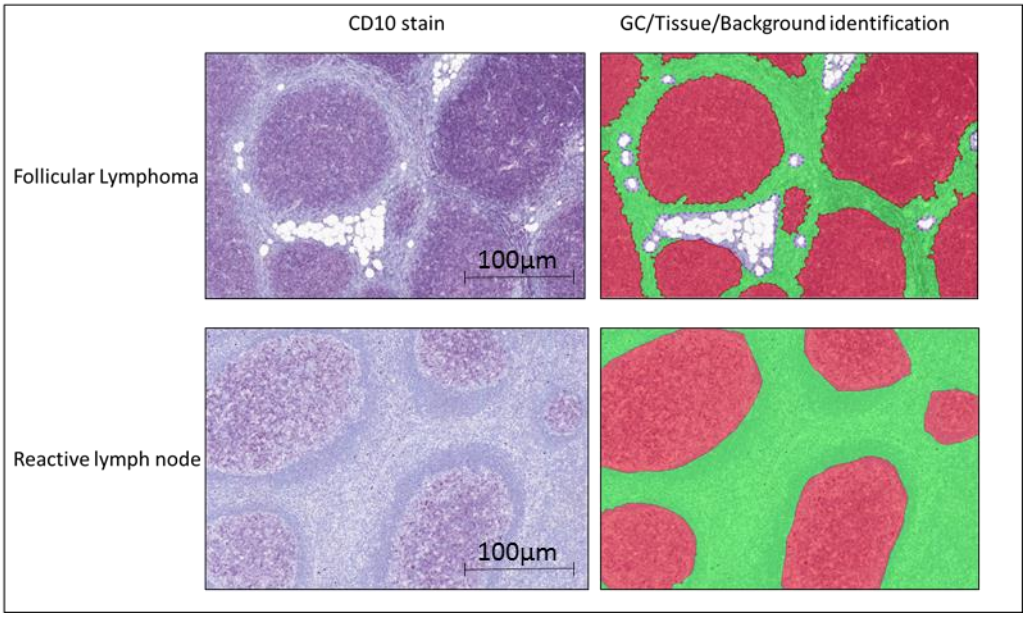


Figure 2.15: *Examples of Visiopharm APP identifying distinct regions following training based on CD10 staining.* GCs are in red, tissue in green and background outlined in blue, which were discarded from downstream analysis. Magnification for both FL and Reactive LN: x11.5.

The APP was run on all samples and checked manually to ensure correct region identification.

The phenotype of cells in both the GC and tissue were identified using another APP, with cell type biomarkers (CD68, CD21, CD163) used to quantify the cells expressing DC-SIGN and MR (Table 2.7). The threshold values for each stain were individually optimised to take into account the different staining intensities of each antibody.

DC-SIGN+ cells		**MR+ cells**	
GC	*Tissue*	*GC*	*Tissue*
CD163+CD68+		CD163+CD68+	
CD163+CD68-		CD163+CD68-	
CD163-CD68+		CD163-CD68+	
CD163-CD68-		CD163-CD68-	
CD21+		CD21+	
MR+		DC-SIGN+	

Table 2.7: *Cell types defined by Visiopharm APP analysed for DC-SIGN and MR positivity.*

The APP analyses staining in the six channels (representing the six stains) and determines whether a specific cell is either unstained, single stained, or multi-stained for several cell markers through multiplexing the individual stains for each tissue section following the alignment described above. Figure 2.16 shows identification of single stained CD68+ and CD21+ cells whereas Figure 2.17 highlights identification of MR and DC-SIGN positivity in CD68+ and CD163+ cells. The APP identifies positive stains while excluding weaker, less specific stains across all six channels.

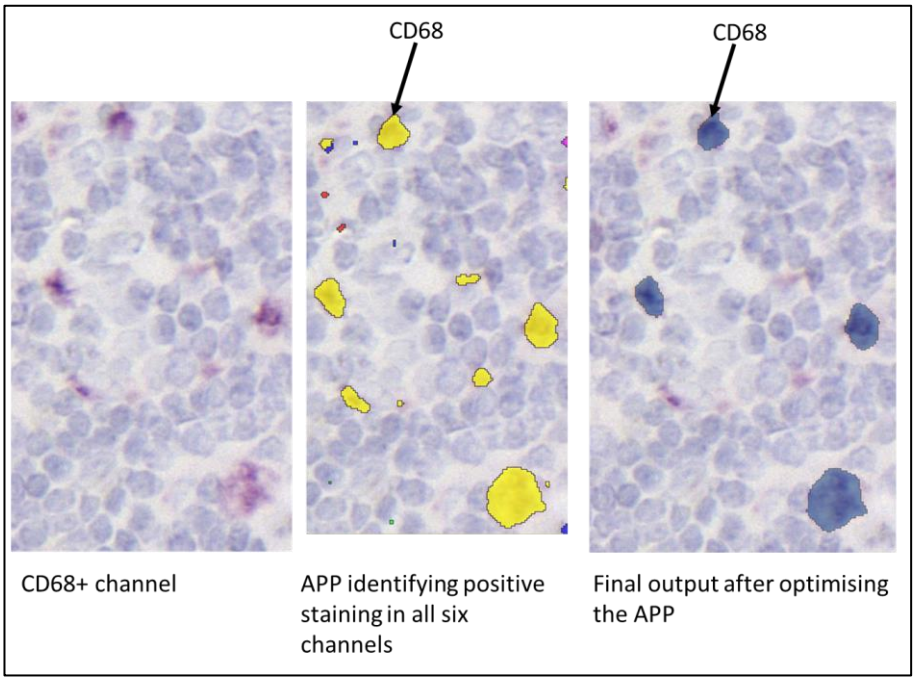


Figure 2.16: *Identification of CD68 and CD21 positive cells through staining intensity*. All six channels were analysed and depending on whether the staining threshold value is reached, the APP will identify a cell positive or negative for a single stain that will be compared to other markers in the section, determining whether a cell expresses one or multiple markers. For these cells, only CD68 or CD21 were identified as positive markers. Magnification of top panel images: x85.45. Magnification of bottom panel images: x19.9.

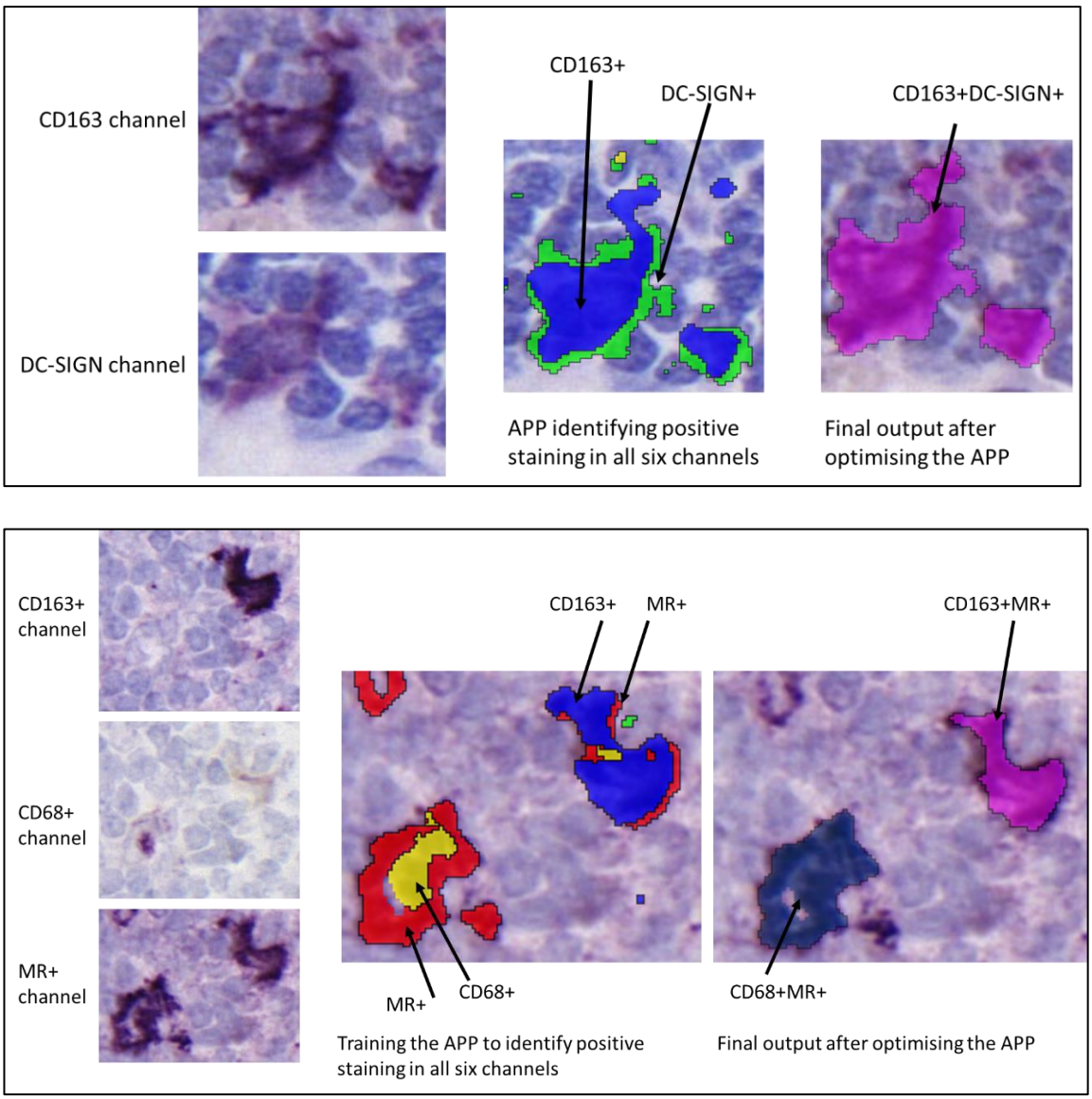


Figure 2.17: *Identification of MR and DC-SIGN positivity in CD163+ and CD68+ cells*. Magnification of top panel images: x212.62. Magnification of bottom panel images: x102.54.

Quantitative output included mean intensity of staining of positive cells, their actual numbers, and their location within the tissue (e.g. GC or non-GC). Advantages of automated scoring compared to manual pathologist scoring includes avoiding bias in staining interpretation (Varghese et al., 2014, Choudhury et al., 2010) along with its high throughput capability which would otherwise be time-consuming by manual analysis.

Chapter 3: ACQUIRED N-GLYCOSYLATION MOTIFS AND THEIR ROLE IN FL EVOLUTION

3.1 Introduction

As mentioned previously, FL is a clinically heterogeneous disease with some patients responding extremely well to first-line therapy whilst others progress or transform within two years of treatment. The course of disease is characterised by several remitting-relapse cycles, with patients becoming increasingly refractory to treatment. Transformation of FL to a high grade, aggressive lymphoma is a critical event in disease evolution, commonly associated with poorer clinical outcomes. Genomic analysis of temporal FL biopsies including cytogenetic and CNV analysis has indicated that disease progression is not a step-wise process with the accumulation of mutations in major clones, but instead occurs through a divergent pathway, founding from a putative CPC population (Fitzgibbon et al., 2007, Johnson et al., 2008, Martinez-Climent et al., 2003). Shared mutations between disease episodes is indicative of mutations harboured within the CPC, and based on the degree of genetic semblance between episodes, the CPC for a patient can either be 'rich' or 'sparse' (Okosun et al., 2014). Sparse indicates an early divergence of event specific clones from the CPC, with their evolution occurring through pathways that are more independent. As this CPC reservoir pool is believed to give rise to successive disease events, uncovering key pathogenic events within this population may lead to the development of therapeutic interventions that abrogate the remitting-relapse cycle of disease and improve the outcome for patients. Recurrent mutations that are clonally maintained over temporal FL/tFL samples occur in histone modifying genes, including CREBBP and KMT2D, indicating epigenetic 'addiction' as a feature of both the 'rich' and 'sparse' CPC (Okosun et al., 2014).

However, to our knowledge, no studies have explored the microenvironmental interactions of the CPC, despite FL cells being highly dependent on cross-talk with immune cells of the GC (Ame-Thomas and Tarte, 2014) . A basic demonstration of this is seen by the difficulty of growing FL cells *in vitro* without stimulation through CD40 (Ghia

et al., 1998) engagement, a main signalling pathway of B and T cell interaction. The selective retention of the BCR throughout disease highlights another critical survival-signalling pathway used by FL cells.

Newly discovered replacement mutations caused by SHM introduces amino acid sequence motifs into the IGHV of the BCR, consisting of Asparagine (Asn)-X-Serine/Threonine (S/T), where X can be any amino acid except proline. These N-glycosylation (N-gly) motifs are found in ~90% and 73.5% of surface IgM and IgG positive FL cases, respectively (McCann et al., 2006, Kuppers and Stevenson, 2018). These motifs occur in the heavy chain IGV of only 9% of healthy B cells (Alcoceba et al., 2012), indicative of a pathogenic function. Oligosaccharides of the high mannose type are added to motifs and are retained on the Ig, despite being an unusual feature of cell surface proteins (McCann et al., 2008). Oligomannoses can interact with calcium-dependent lectins associated with cells of the microenvironment, activating the BCR signalling pathway and likely contributes to the survival and retention of tumour cells in the GC (Coelho et al., 2010). This retention may enable the accumulation of recurrent mutations characteristic of FL, including those involved in epigenetic modification.

The behaviour of N-gly motifs during disease evolution and progression has been investigated through *IGHV* cloning in a number of FL cases. Studies have indicated a conservation of acquired N-gly sites within identified molecular clones (McCann et al., 2006). However, clone numbers have been limited in such studies and do not truly reflect the extent of intraclonal diversity that exists in disease as a result of ongoing SHM. Furthermore, as the analysis has been restricted to a single disease event, behaviour over time has not been addressed and therefore no conclusions regarding their role in the putative CPC has been determined. To address this requires comprehensive IGV analysis of the clonal repertoire taken from temporal biopsies, ranging from a relatively early time point in disease manifestation (e.g. diagnosis) to a time point at which the disease has become genetically and clinically distinct (e.g. relapse and transformation). As SHM continues during disease progression and transformation, the stepwise process can be visualised through lineage trees rooted to a putative non-malignant germline IGV sequence, making them an important tool in B cell evolutionary studies.

3.2 Objectives

By implementing an NGS approach, this study set out to identify and catalogue N-gly motifs in the *IGHV* gene of a patient's lymphoma as it progressed from an indolent form to an aggressive form of the disease, tFL. A longitudinal profiling strategy enabled the delineation of the clonal *IGHV* repertoire throughout a patient's disease course, giving insight into the importance of motifs during disease evolution, expanding our views on the ontogeny and pathogenesis of FL. The inclusion of a clinically and biologically variable patient cohort ensured findings were representative of the heterogeneous disease.

3.3 Methods

N.B The term 'clone' mentioned throughout this chapter is solely based on *IGHV* sequencing data and does not reflect mutational status/mutational burden. The 'major clone' refers to the *IGHV* sequence identified through Sanger sequencing of the homoduplex. We can infer this is representative of the major clone as successful sequencing indicates a high copy number of the sequence in the sample which has outcompeted minor clone sequences. In NGS data, the major clone was identified by having the highest count number.

3.3.1 Patient samples

Genomic DNA for all FL samples used in this study were readily available from our department (Haemato-Oncology). This included 30 diagnostic FL samples, twelve paired FL-tFL samples and three additional paired FL-tFL samples which were used in NGS. DNA was previously extracted from biopsies that were histologically reviewed to confirm the diagnosis. *IGHV* sequencing files for three temporal FL samples were available from the European Nucleotide Archive (ENA) using the accession numbers PRJEB9334 and PRJNA240336 (Jiang et al., 2014, Carlotti et al., 2015) to validate findings.

3.3.2 BCL2-IGH breakpoint analysis

Clonality between paired samples was confirmed by identical BCL2-IGH breakpoints, which is detailed in Section 2.15.

3.3.3 *IGHV* amplification

Amplification of *IGHV* was performed by multiplex PCR using Biomed consensus primers targeting the Framework 1 and J_H regions, as detailed in Chapter 2. A high fidelity polymerase with a proofreading 3'-5' exonuclease domain was used to minimise polymerase reading errors thus ensuring that observed SNPs were generated by the SHM process. Details of PCR reaction mixture and cycling conditions are detailed in Section 2.4.

3.3.4 Heteroduplex analysis

To distinguish monoclonal products from polyclonal products, PCR products underwent heteroduplex analysis. The PCR product was denatured and renatured at 95^{o}C and 50^{o}C, respectively. Products were run on a polyacrylamide gel to distinguish homoduplexes from heteroduplexes, with full details found in Section 2.5. Homoduplexes representing the major tumour clones were extracted and underwent further PCR amplification. The PCR product was then purified before undergoing sequencing or cloning.

3.3.5 Cloning

To identify tumour subclones, PCR products were cloned using the TOPO TA Cloning® system, described in detail in Section 2.16. Up to ten colonies were selected for PCR amplification and Sanger sequencing.

3.3.6 Sanger sequencing

Purified PCR product was sent to Eurofins Genomics Sanger sequencing services (Germany) for bidirectional capillary electrophoresis using the ABI PRISM 3730XL Genetic Analyser, as detailed in Section 2.11. PCR products were sequenced in both the 5' and 3' direction to ensure accurate sequencing of the entire V gene.

3.3.7 VDJ assignment and SHM analysis

The *IGHV* sequences produced by Sanger sequencing were entered into the IMGT/V-QUEST programme available at http://www.imgt.org, which compares sequences to all known germline immunoglobulins, identifying the V, D and J gene segments of closest homology. SHM sites were identified through discordance between nucleotides of experimental and germline sequences. The sites of SHM were compared between the major clones of each paired FL-tFL sample in addition to the multiple disease-specific subclones derived from cloning. SHM events gave rise to either synonymous (e.g. no change to the amino acid encoded by the affected codon) or non-synonymous (e.g. change to the amino acid sequence) mutations which were identified by the IMGT/V-QUEST programme. Full details of IMGT/V-QUEST are found in Section 2.12. For sequencing data generated from next-generation sequencing platforms, the IMGT-HighV-QUEST program was used (Section 2.19).

3.3.8 N-gly motif identification

The NetN-glyc 1.0 online server (www.cbs.dtu.dk/services/NetNGlyc/) was used to identify N-gly motifs in *IGHV* sequences, as detailed in Section 2.13.

3.3.9 Amplicon sequencing

Re-amplified homoduplex PCR products were sent to Genewiz (United Kingdom) for paired-end 2x250bp sequencing using the Illumina Miseq platform, as detailed in

Section 2.17. The full pipeline regarding processing and subclone identification based on *IGHV* reads are also detailed in Section 2.18 and 2.19.

3.3.10 Lineage tree

To visualise clonal expansion and the relationship between individual subclones, lineage trees were constructed using the IgTree program algorithm (Barak et al., 2008). Full details of this process can be found in Section 2.24.

3.4 Results

3.4.1 N-gly detection in a diagnostic FL cohort

Diagnostic FL samples were used to optimise *IGHV* amplification and validate previous findings regarding N-gly motif frequency in disease.

Homoduplex *IGHV* bands located at the 300-400bp region were identified for ten diagnostic samples (Figure 3.1). The variation in migration through the gel was indicative of unique *VDJ* rearrangements between cases.

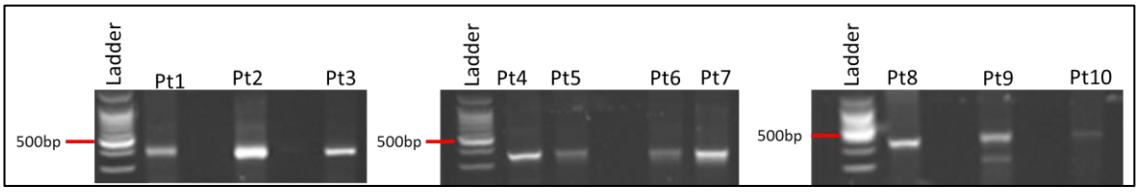


Figure 3.1: *IGHV bands visualised on agarose gel following amplification of homoduplexes*. Homoduplexes were amplified by FR1 and J_H consensus primers, with primer sequences listed in Figure 2.1. Bands were extracted followed by DNA gel extraction before being sequenced. The faint lower band for Patient 9 was also sequenced and had no results, suggesting an artefact.

Patient	Germline IGHV	Homology (%)	No. of sites	Region	aa Motif
1	V4-59, J4, D5-18	90.75	1	CDR2	NIS
2	V4-39, J5, D2-21	94.32	2	CDR2, FR3	NIS, NAS
3	V4-61, J6, D4-17	92.44	1	CDR1	NWT
4	V3-48, J4, D4-11	89.13	2	CDR1, CDR2	NMT, NIS
5	V3-48, J6, D5-18	85	3	CDR1, CDR2, FR3	NMT, NIT, NIS
6	V3-48, J4, D3-10	86.11	3	CDR1, CDR2, CDR3	NMT, NIS, NYS
7	V4-59, J5, D5-24	93.64	1	CDR2	NVS
8	V3-48, J4, D2-2	85.42	1	CDR2	NIT
9	V4-38, J4, D3-10	89.86	1	CDR2	NIS
10	V4-31, J4, D3-10	88.97	0	-	-

Table 3.1: *Incidence of novel N-glycosylation sites in diagnostic FL samples*. CDR-complementarity determining regions, FR-framework region. Homology indicates degree of similarity between the experimental sequence and the germline of closest homology.

In 90% of cases, motifs were identified giving support to the high prevalence seen in other studies (Zhu et al., 2002) (Table 3.1). Patients either had one or multiple sites. As none of the motifs were found in the corresponding germline sequences of the *IGHV* gene and arose from base substitution (Figure 3.2), SHM is the cause for acquisition. Without a germline DNA sample, it is difficult to determine the nature of the CDR3-associated N-gly motif of patient 6 as theoretically, nucleotides could have been added to junctional regions during the somatic recombination. However, the lack of motifs in CDR3 regions of normal somatically mutated B cells is strong evidence that motifs are acquired through SHM and are positively selected in FL.

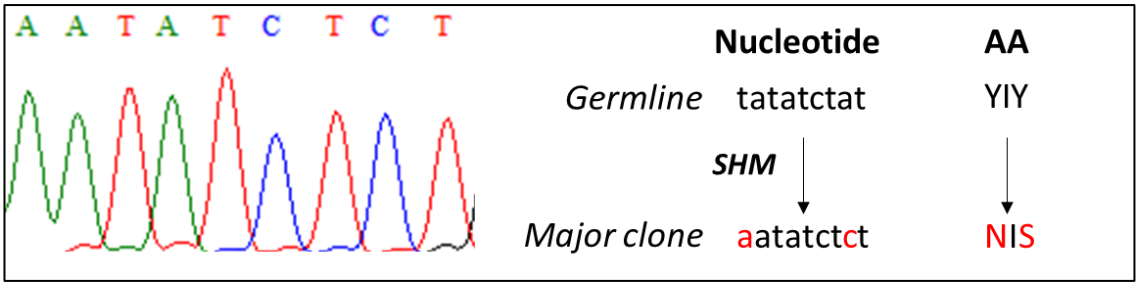


Figure 3.2: *Acquired N-gly motif of patient 1.* The electropherogram identifies the sequence encoding the N-gly motif. The germline nucleotide sequence of *V4-59* was compared to the experimental sequence, highlighting that point mutations lead to motif acquirement (YIY>NIS), as indicated by red letters. AA-amino acid.

Analysis of the acquired asparagine residue of the N-gly motif revealed that in all cases, the amino acid was dissimilar or highly dissimilar to its germline counterpart (Table 3.2), based on three values; hydropathy, volume and physiochemical properties.

Patient	SHM hotspot (Y/N)	Comparison between asparagine and G.L aa
1	N	Dissimilar
2	N, N	Very dissimilar, dissimilar
3	N	Very dissimilar, dissimilar
4	Y, N	Very dissimilar, very dissimilar
5	N, N, Y	Dissimilar, dissimilar, very dissimilar
6	Y, N, N/A	Dissimilar, very dissimilar, N/A
7	N	Dissimilar
8	N	Dissimilar
9	N	Dissimilar

Table 3.2*: Location of N-gly motifs in relation to known SHM hotspots*. Known SHM consist of rgyw (a/g)g(c/t)(a/t) or wrcy (a/t)(a/g)c(c/t) sequences. The similarity of the acquired asparagine compared to the original germline amino acid is based on three values; hydropathy, volume and physiochemical property of the amino acid. Dissimilar refers to two out of three values being different, and very dissimilar refers to all three values being different. The CDR3 located motif was not able to be analysed due to a lack of a comparative germline sequence. N/A-not available.

An extension cohort of twenty diagnostic samples (Supplementary Table 1.1) revealed that the majority of motifs occurred within the hypervariable regions, with 54% occurring in the CDR2 (Figure 3.3a). Although the framework regions maintain the

structure of the protein and are therefore more conserved, they are still prone to SHM, especially during prolonged GC expansion, as seen by the acquirement of motifs in FR3 in four patients.

Out of 30 samples, only nine *V* gene subgroups were used, highlighting an over-representation of *V* genes in disease (Figure 3.3b). The most recurrently used genes included *V3-48*, *V3-11* and *V3-23*. Interestingly, the germline amino acid sequence of these genes display motif 'starter sequences'. These sequences can generate motifs with either one or two nucleotide substitutions in the first amino acid. For example, the CDR2 region of *V3-48* and *V3-11*, contains a 'YIS' amino acid sequence which is frequently substituted to 'NIS'. *V3-23* contains YIS, AIS and TIS in the CDR2, CDR2 and FR3 regions, respectively. As these starter sequences frequently give rise to motif sites in our cohort, this could explain the *V* gene bias observed in the disease.

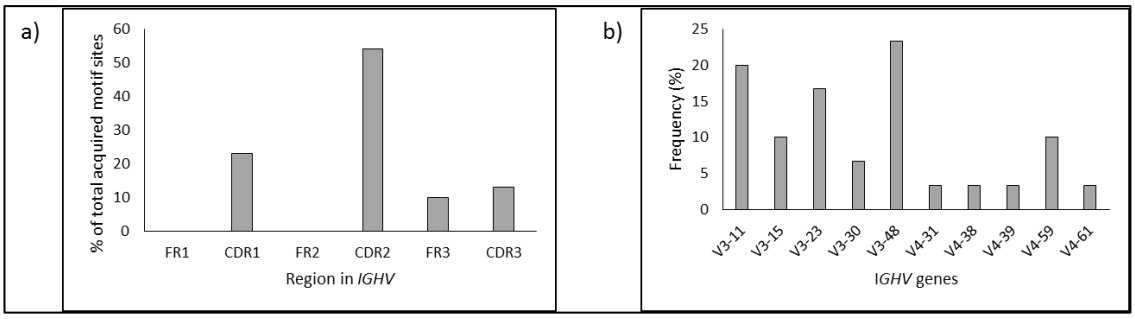


Figure 3.3: *Analysis of IGHV gene in 30 diagnostic FL samples*. a) Distribution of N-gly sites in variable region. Values are expressed for each region as a percentage of the total number of sites (n=39). b) Usage frequency of V gene subgroups. Values for each subgroup is expressed as a percentage of the total number of patients (n=30).

3.4.2 Paired FL- tFL samples

During the evolution of many malignancies, there is remodelling of the clonal repertoire due to selection pressures (e.g. therapy), leading to the outgrowth of clones which have a genetic advantage that progress the disease (McGranahan and Swanton, 2017). This is also observed in FL with the ongoing SHM process enabling us to track which subclones (based on *IGHV* sequences) gain dominance (Horn et al., 2016). Determining whether

motifs are conserved in the major clones of successive disease events despite an assumed change in their mutational profile can indicate their importance during progression.

Genomic DNA from twelve paired FL-tFL samples was readily available. Clinical information regarding these patients is given below (Table 3.3). Patient 15 had three temporal samples. The median age of diagnosis was 47.5 years with a median time to transformation of 4.6 years, ranging from 0.6 years to 14.6 years. Therapy information was available for some patients.

Pt No.	Gender	Age at diagnosis	Histological grade at diagnosis	Sample types	Time to transformation from diagnosis (months)	No of lines of therapy before transformation	Line of therapy
11	F	45	1	Diagnosis/Transformation	40	1	FL(diagnosis)-Radiotherapy (Jul 99) 1st relapse- Chlorambucil Transformation- CHOP, then LD-56 trial, then Rituximab (Nov 02)
12	F	30	1	FL relapse 7 /Transformation	174.9	13	NA
13	M	29	4	Diagnosis/Transformation	7	0	FL (diagnosis)-Expectant management (Jul 11)
							Transformation-R-CHOP (Feb 12)
14	M	45	2	FL relapse 1 /Transformation	44.9	7	1st relapse- Cytarabine + Etoposide, followed by BEAM autograft (Jul 99)
15	M	57	3	Diagnosis/4th relapse/ /Transformation	81	3	FL(diagnosis)-Chlorambucil, then CHOP (April 01) 4th relapse-Velcade + Rituximab Transformation-R-CHOP + RIC-SCT (Jan 08)
16	M	67	2	Progression/Transformation	35.7	1	NA
17	F	42	1	Diagnosis/Transformation	39.5	0	NA
18	F	32	NA	FL Relapse 2 /Transformation	72.9	4	NA
19	M	61	2	FL Relapse 2 /Transformation	22.8	3	NA
20	F	47	1	Progression/Transformation	94.5	7	NA
21	F	70	2	Progression/Transformation	27.5	0	NA
22	M	45	1	Diagnosis/Transformation	18.2	2	NA

Table 3.3: *Clinical characteristics of patients included in paired sample analysis*. Pt-patient, CHOP-cyclophosphamide, doxorubicin, vincristine, prednisolone; R-rituximab; RIC-SCT-reduced intensity conditioning stem cell transplantation.

To ensure transformation was not *de novo* DLBCL, clonality between paired samples was verified through identical BCL2 breakpoints (Figure 3.4) and *VDJ* rearrangements.

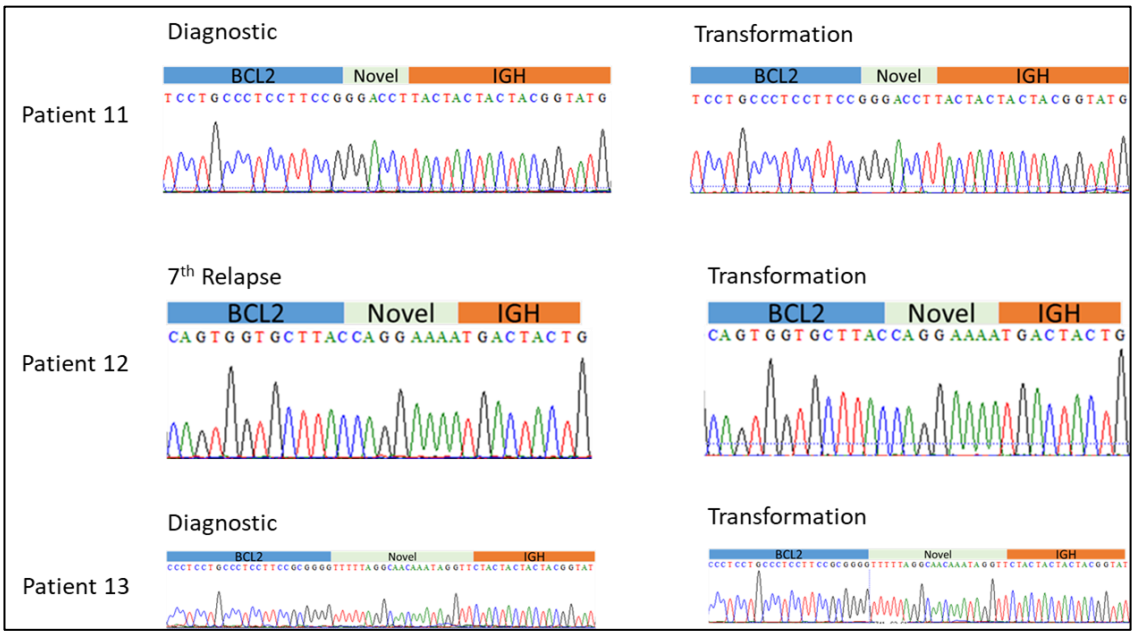


Figure 3.4: *Shared BCL2-IGH breakpoints for 3 selected FL-tFL paired cases*. Genomic DNA taken from 3 patients at different disease event time points was used to amplify the BCL2-IGH breakpoint, using MBR and JH primers, with their sequences listed in Table 2.1. The breakpoints were identified by entering sequences into BLAST and comparing to reference sequences of BCL2 (Accession No. NC_000018.10) and IGH (Accession No. NC_000014.9). Shared breakpoints between paired samples indicates clonality.

3.4.3 SHM in paired FL-tFL

SHM patterns in the *V* gene region were also assessed between pairs, with all samples having a high number of shared mutations, further confirming a common clonal origin (Figure 3.5). Additional unique SHM sites also occurred in both FL and tFL biopsies, suggesting a divergent evolution from an earlier founder clone. This is supported by previous studies of SHM in FL progression (Carlotti et al., 2009). SHM resulted in several shared and unshared non-synonymous mutations that are highlighted in patients 11-13 (Figure 3.6).

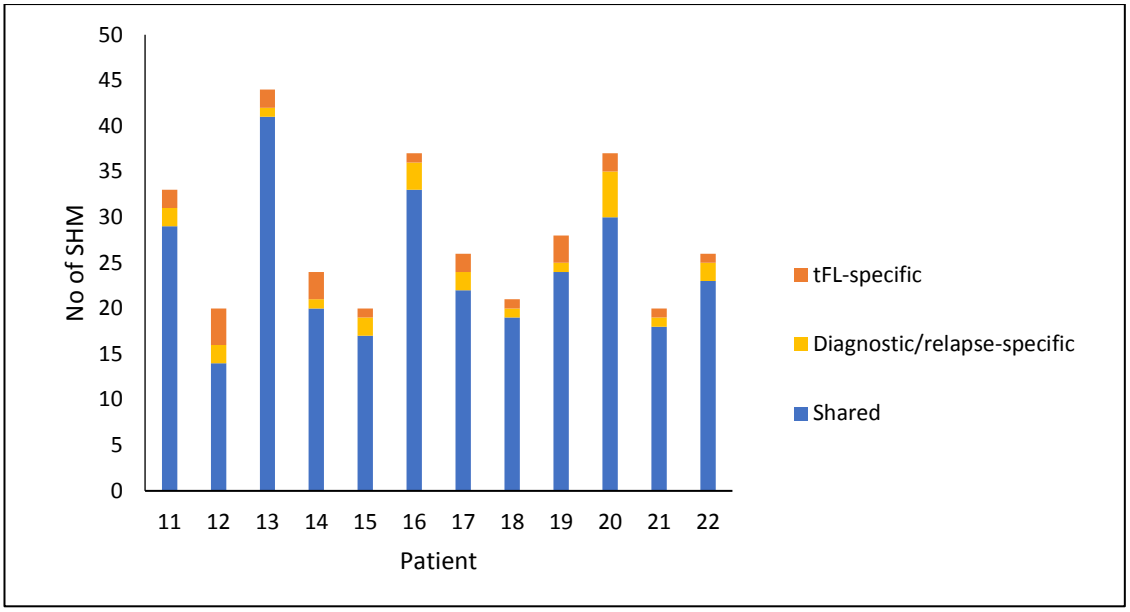


Figure 3.5: *SHM within the heavy chain V gene of paired FL-tFL samples*. SHM is defined as a variation in nucleotide sequence of the major clone V region from the germline V region sequence of closest homology. Shared mutations are identified (blue), in addition to disease-specific mutations occurring at unique sites within the V gene. For patient 15, the 4th relapse sample was omitted from the graph to compare SHM between the most temporally distant samples.

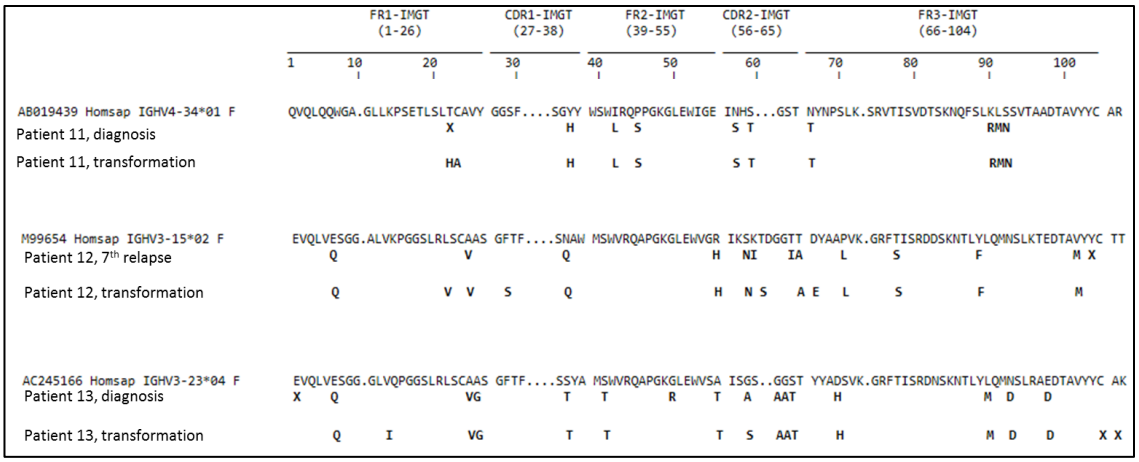


Figure 3.6: *Comparison of the V gene amino acid sequence for three FL-tFL paired cases*. The top lines in the three cases represent the germline amino acid sequence of the *V* gene with the closest homology, as identified by IMGT-V-QUEST. Letters in bold represent distinct amino acids to the germline sequence in each disease event, a result of non-synonymous mutations introduced by SHM. Comparison of amino acids between paired samples shows shared and distinct SHM within each disease event.

3.4.4 N-glycosylation motifs during disease progression

All patients with the exception of patient 20, contained N-glycosylation motifs across the variable region (n=11) (Table 3.4). Patients either had one or multiple N-gly motifs. The majority of sites occurred in the CDRs with additional sites in the FR2 and FR3. Non-synonymous mutations lead to the acquisition of the asparagine residue in all patient samples, as revealed by comparison to the germline sequence. The serine/threonine amino acid of the motif was either a feature of the germline nucleotide sequence or gained through synonymous mutation or non-synonymous mutation. The extent of SHM within the acquired motif site is indicated in Table 3.5.

Patient	FL Disease Event	*IGHV* gene	No. of acquired sites	Region	aa Motif
11	FL	V4-34	1	CDR3	NST
	t-FL(DLBCL)	V4-34	1	CDR3	NST
12	FL7	V3-15	1	CDR2	NIT
	t-FL(DLBCL)	V3-15	1	CDR2	NKS
13	FL	V3-23	1	CDR3	NFS
	t-FL(DLBCL)	V3-23	1	CDR3	NFS
14	FL	V4-59	1	CDR2	NVT
	t-FL(DLBCL)	V4-59	1	CDR2	NVT
15	FL	V3-7	4	CDR1, CDR1, CDR2, FR3	NFS, NYS, NET, NLS
	FL4	V3-7	4	CDR1, CDR1, CDR2, FR3	NFS, NYS, NET, NLS
	t-FL(DLBCL)	V3-7	4	CDR1, CDR1, CDR2, FR3	NFS, NYS, NET, NLS
16	FL1	V3-48	1	CDR3	NLS
	t-FL(DLBCL)	V3-48	1	CDR3	NLS
17	FL	V3-48	1	CDR2	NIT
	t-FL(DLBCL)	V3-48	1	CDR2	NIT
18	FL2	V3-48	1	FR2	NMT
	t-FL(DLBCL)	V3-48	1	FR2	NMT
19	FL2	V3-23	4	CDR1, CDR2, CDR2,CDR3	NIT, NIS, NKT, NCT
	t-FL(DLBCL)	V3-23	4	CDR1, CDR2, CDR2,CDR3	NIT, NIS, NKT, NCT
20	FL	V3-23	0	-	-
	t-FL(DLBCL)	V3-23	0	-	-
21	FL	V3-11	1	CDR1	NFS
	t-FL(DLBCL)	V3-11	1	CDR1	NFS
22	FL	V3-30	1	CDR3	NFS
	t-FL(DLBCL)	V3-30	1	CDR3	NVS

Table 3.4: *N-glycosylation sites acquired through SHM in the VH gene of paired FL/t-FL cases.* Identical V gene usage between the dominant clones in sequential samples highlights the clonal relationship between disease events. FL-1 indicates a first relapse. t-FL indicates a transformation. N-gly motifs which differ in amino acid composition between temporal samples are indicated in red.

Patient	FL Disease Event	AA Change	Bp mutation[A]		
			1	2	3
12	FL7	SKT>NIT	AAC	ATA	ACT
	t-FL(DLBCL)	SKT>NKS	AAC	AAA	AGT
14	FL	YIY>NVS	AAC	GTC	TCT
	t-FL(DLBCL)	YIY>NVS	AAC	GTC	TCT
15	FL	TFS>NFS, SYW>NYS, SEK>NET, SLR>NLS	AAT, AAT, AAT, AAT	TTT, TAT, GAG, TTG	AGT, TCC, ACG, AGC
	FL4	TFS>NFS, SYW>NYS, SEK>NET, SLR>NLS	AAT, AAT, AAT, AAT	TTT, TAT, GAG, TTG	AGT, TCC, ACG, AGC
	t-FL(DLBCL)	TFS>NFS, SYW>NYS, SEK>NET, SLR>NLS	AAT, AAT, AAT, AAT	TTT, TAT, GAG, TTG	AGT, TCC, ACG, AGC
17	FL	IST>NIT	AAT	ATT	TAC
	t-FL(DLBCL)	IST>NIT	AAT	ATT	TAC
18	FL2	GKG>NMT	AAT	ATG	ACT
	t-FL(DLBCL)	GKG>NMT	AAT	ATG	ACT
19	FL2	TFS>NIT, AIS>NIS, ATT>NKT	AAC, AAT, AAT	AAT, AAT, AAG	ACC, AGC, ACG
	t-FL(DLBCL)	TFS>NIT, AIS>NIS, ATT>NKT	AAC, AAT, AAT	AAT, AAT, AAG	ACC, AGC, ACG

Table 3.5: *Analysis of somatic mutations giving rise to N-gly sites*. Red letters indicate a nucleotide replacement mutation.[A] refers to amino acid position 1, 2 and 3 of N-gly motif. Patients 11, 13, 16, 21 and the fourth motif site in patient 19 were excluded from analysis due to motifs being located in the CDR3 region which does not have comparative germline sequence to determine somatic mutations at motif site.

For patients 14, 15, 17, 18, 19 and 21, motif sites occurred in the FR1 to FR3 region and both site location and nucleotide sequence were conserved in progression and transformed samples (Figure 3.7). This conservation was not dependent on the number of motifs present in the earliest event, as seen in patients 15 and 19 which conserved all four sites. This conservation highlights that for these patients, motifs had arisen from shared SHM in a common ancestral cell that gives rise to the successive disease events. This common ancestral cell does not appear to be influenced by therapy as highlighted in patient 15 who had undergone several lines of therapy between biopsies (Table 3.3).

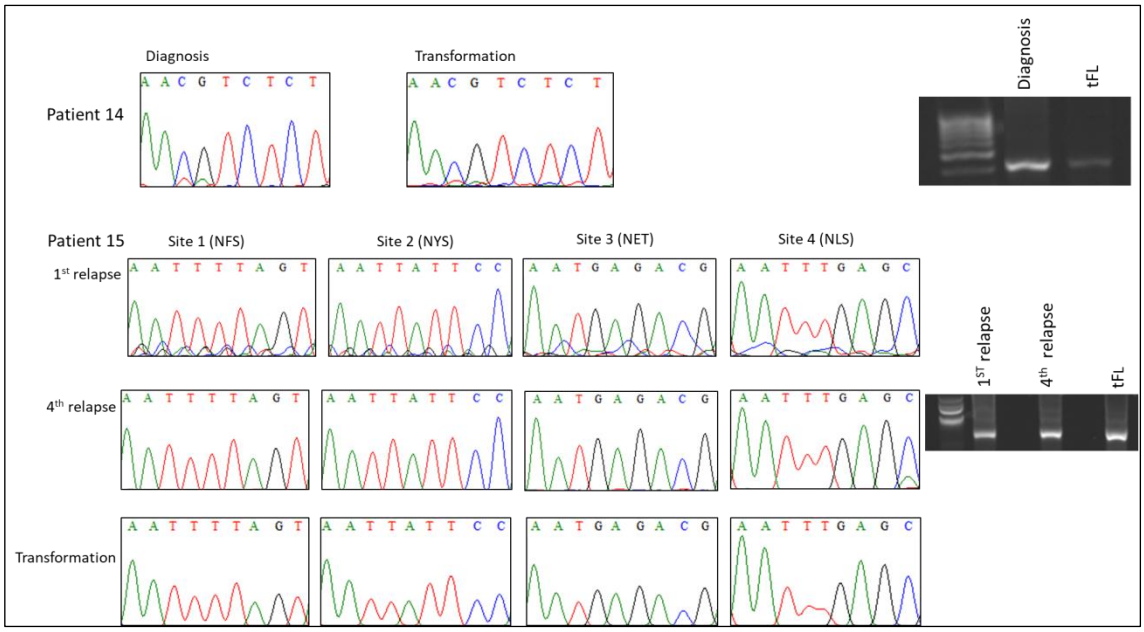


Figure 3.7: *N-gly motif sites in two paired FL-tFL patient samples*. Identical nucleotides between sequential samples indicate motifs have been acquired through shared SHM events in a common ancestral cell. IGHV PCR products used to sequence are shown on the right.

N-gly motifs were also present in the CDR3 region of patients, 11, 13, 16, 19 (Figure 3.8). Although there is no germline CDR3 sequence to determine the origin of these motifs, their absence in the CDR3 region of healthy B cells is strong evidence that they were acquired through SHM, and similar to motifs in the FR1 to FR3 region, their conservation in temporal samples suggests they were acquired in a shared ancestral cell.

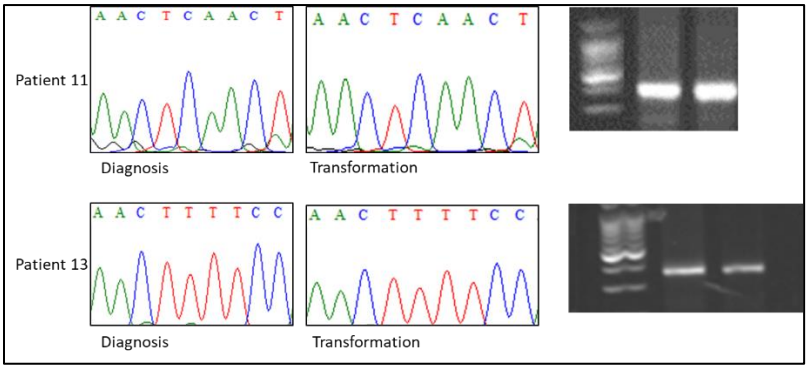


Figure 3.8: *Nucleotide sequences encoding for N-gly motifs in the CDR3 region of paired FL-tFL samples.* Nucleotide composition is maintained as the disease progresses. Corresponding IGHV bands are indicated on the right.

3.4.5 SHM variation within N-gly motif sequence

For patients 12 and 22, we observe a similar conservation of motif sites. Yet a distinction from the other patients lies in the difference in amino acid sequences between temporal samples (Table 3.4). For patient 12, the relapse and transformed motif amino acid sequence located in CDR2 were NIT and NKS, respectively. This suggests that the motifs seen in the different disease events are a result of unshared SHM. However, analysis of the codon giving rise to the asparagine residue of the motif reveals an identical nucleotide composition (aac) between the two samples (Figure 3.9), suggesting that this shared SHM event preceded the unshared mutations seen in the second and third codon as SHM is a stepwise process. If we relate this back to the germline (*V3-15* gene) amino acid sequence which consists of SKT (serine, lysine and threonine), the shared SHM event would give rise to the N-gly motif, NKT. This indicates that temporal samples of patient 12 likely arose from a shared ancestral cell that displayed this motif. A divergence in evolution as evidenced by unshared SHM between events, then lead to the distinct motif compositions. As the unshared SHM maintained the motif site, we can infer a positive selection of motif containing clones during disease progression.

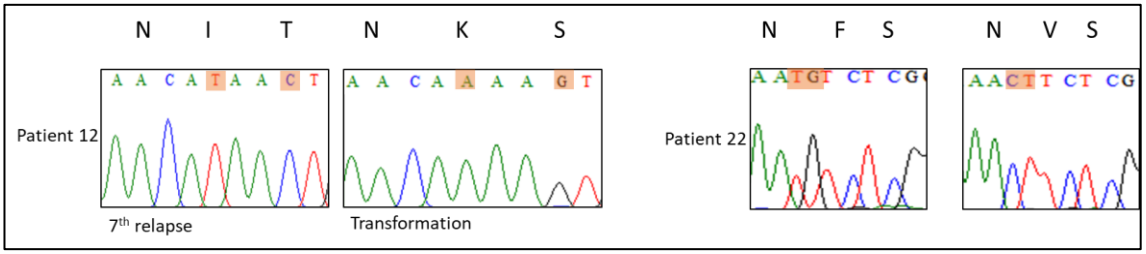


Figure 3.9: *Paired FL-tFL samples that have distinct SHM in the nucleotide sequences encoding for N-gly motifs.* For patient 12, changes to the second and third codon in the transformed samples results in amino acid changes. For patient 22, a change to the second codon also result in an amino acid change. Despite changes to the first codon representing asparagine in patient 22, the amino acid is retained, maintaining the motif site. Differences in nucleotides between temporal samples are highlighted in orange.

For patient 22, the CDR3-based motif in the diagnostic and transformed samples was NFS and NVS, respectively. Unlike patient 12, the codon giving rise to the asparagine

residue differed between samples; aac encoded for asparagine in the diagnostic sample and aat encoded for asparagine in the transformed sample (Figure 3.9). This nonsynonymous mutation could be due to ongoing SHM in the first codon, which is a valid viewpoint given the largely random nature of the SHM process. Yet the selection of a major clone which conserves the motif site, indicates that motif positive clones have an advantage during disease progression.

3.4.6 Fate of germline encoded motifs compared to acquired motifs

For patient 11, the germline *V4-34* gene contains a naturally occurring motif site in the CDR2 region (NHS). However in both samples, a shared SHM (a>g) resulted in substitution of the asparagine residue to a serine, eliminating the motif from a shared ancestral cell (Figure 3.10). Acquisition of a new motif site in the CDR3 region was gained through shared SHM. The loss of the *V4-34* natural motif site is seen in other studies (Zhu et al., 2002) (McCann et al., 2008) which suggest an irrelevance of sites to disease progression unlike those gained by SHM. This adds support to the claim that natural sites appear irrelevant to disease progression unlike those gained by SHM. This could be due to the differential glycosylation fates of natural versus acquired sites, with natural sites not found to be glycosylated.

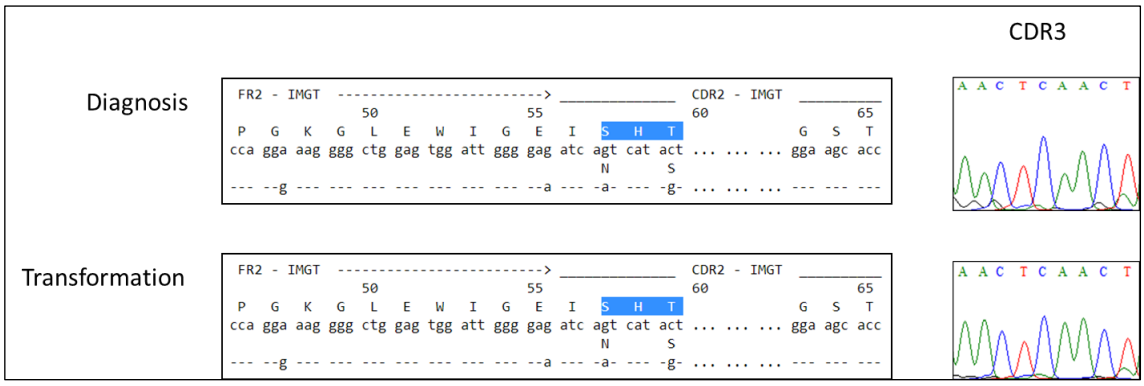


Figure 3.10: *Analysis of germline and acquired N-gly motifs for patient 11's temporal samples. V4-34* contains a naturally occurring N-gly motif in the CDR2 region (NHS). IMGT analysis of experimental sequences reveals loss of this site and replacement by SHT (indicated in blue highlight). Identical nucleotide sequences highlight that loss was a result of shared SHM in a common ancestor. The acquired site was gained through shared SHM.

3.4.7 N-gly motifs in intraclonal variants

Although the major clone makes up the bulk of the tumour, heterogeneity of disease is attributed by the unique subclones present as a result of ongoing SHM which may have distinctive mutational profiles. If a minor subclone gains an advantageous mutation, it can gain dominance, enabling the disease to evolve. This flux in clonal dynamics highlights the importance of evaluating subclones when considering disease progression.

To determine whether motifs are a feature of intraconal variants, heteroduplex derived *IGHV* products from five paired FL-tFL samples underwent cloning. Up to 10 colonies were selected at random and subsequently sequenced following PCR amplification (Figure 3.11).

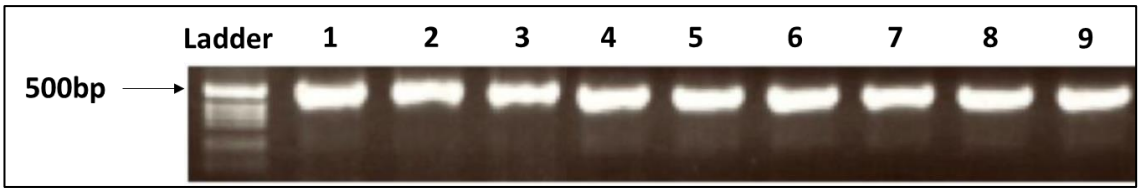


Figure 3.11: *IGHV bands derived from cloning.* Nine white colonies were picked at random from the agar plate of patient 14, FL1 disease event, following the cloning method described in section 2.16. Each colony was amplified by PCR using the M13 forward and reverse primers (primer sequences detailed in section 2.16). Each amplified colony is represented in this figure by a single band.

The majority of amplified colonies contained an identical sequence to the major clone and were therefore discarded. Unique subclones were identified by having distinct sequences of ≥1 nucleotide between each other and the major clone. The number of subclones identified from the total number of colonies sequenced in the paired samples is given in Table 3.6. Subclones had a similar SHM rate to that of the major clone, with many shared SHMs. Despite also containing distinct SHMs, N-gly motifs identified in the major clone of the disease event were conserved in all clones (Table 3.6). This indicates motifs are a clonal SHM event and therefore occurred early in evolution.

Patient	FL disease event	Motif in major clone	No. of distinctive clones out of total colonies sequenced	No. of distinctive clones with motif of major clone
11	FL	NST	2/10	2/2
	t-FL(DLBCL)	NST	3/10	3/3
12	FL7	NIT	2/10	2/2
	t-FL(DLBCL)	NKS	4/10	4/4
13	FL	NFS	3/8	3/3
	t-FL(DLBCL)	NFS	4/10	4/4
14	FL1	NVS	3/9	3/3
	t-FL(DLBCL)	NVS	3/10	3/3
15	FL	NYS,NET,NLS	5/10	5/5
	FL4	NYS,NET,NLS	4/10	4/4
	t-FL(DLBCL)	NYS,NET,NLS	3/10	3/3

Table 3.6: *Conservation of N-gly motifs in distinct tumour related subclones*. White colonies selected from the agar plates underwent IGHV amplification and sequencing to determine whether they represented unique subclones from that of the major clone and interrogate their N-gly status. The fourth column represents how many of the colonies sequenced were unique when compared to the major clone. The fifth column represents whether the unique subclones maintained the N-gly site present in the major clone.

3.4.8 Patients undergoing NGS of *IGHV*

Cloning only provides us with a snapshot of the vast clonal repertoire in FL. This is due to the limitations of the technique. Firstly, the number of unique subclones identified largely depends on which colonies are picked. This is because multiple colonies can contain an identical DNA insert which is influenced by the copy number of the original PCR products that are exposed to the plasmid vectors. To reduce this requires picking and sequencing a large number of colonies which can be labourious and time-consuming. Other potential restriction to identifying unique clones includes how much of the transformed E.coli is spread on the agar plate and whether they contain a variety

of unique DNA inserts. To comprehensively analyse the *IGHV* subclonal repertoire, ultra-deep sequencing of the PCR product was required.

Three additional patients with paired FL-tFL samples were selected to undergo NGS sequencing. Samples had previously undergone somatic variant profiling by WGS or WES, which revealed a 'sparse' or 'rich' disease evolution pattern, based on genetic semblance between temporal samples (Okosun et al., 2014). We wanted to investigate whether there was a difference in N-gly dynamics in such mutationally distinct disease events. Patient's 23 & 25 were categorised as 'sparse' and patient 24 categorised as 'rich'. Heteroduplex analysis identified the major clones, with all three patients displaying a *VH3* gene family. Sequencing *IGHV* reads from three additional patients were available, generated from different sequencing platforms and configurations. Patient 26 and 27 sequencing reads were generated from the Roche 454 Life Sciences Genome Sequencer FLX (2x150bp) and patient 28 sequencing reads were generated from the Illumina Miseq platform (2x150bp). The Genewiz bioinformatics service report for patients 23-25 can be found in supplementary Tables 1.2 - 1.4, and supplementary Figures 1.1 and 1.2.

Patient information is given in Table 3.7. The mean age at diagnosis was 49.8 years. All temporal samples excluding patients 27 and 28 were biopsied from the same anatomical site. By having temporal samples derived from distinct anatomical sites, enables us to analyse clonal populations within distinct microenvironments which may confer distinct selection pressures to tumour evolution.

Pt No	Age at diagnosis	Tumour	Biopsy Site	Histological grade	Time from diagnosis years (months)	Overall survival years (months)	Line of therapy
23	N/A	Diagnostic Transformed FL	N/A N/A	N/A N/A	0 1(6)	N/A	N/A
24	40	FL1 FL3 Transformed FL	N/A N/A N/A	N/A N/A N/A	0(5) 2(5) 4	4(2)	Rituximab + BEAM autograft (Jul 97) Bexxar (Jul 99) Methotrexate (Jun 00)
25	48	FL2 FL3 Transformed FL	Left axillary node N/A Left axillary node	N/A N/A N/A	12(3) 15(8) 15(9)	N/A	Expectant management (Aug 00) N/A R-CHOP (Mar 04)
26	43	FL FL2	Left inguinal node	Grade 2 Grade 1	2(8) 4	14(7)	CHOP, Fludarabine, Rituximab, Chlorambucil, Bortezomib Interferon, MethylPrednisolone
27	68	FL FL2	Right axilla Right femoral node	Grade 3A Grade 1	1(4) 1(9)	9(2)	No treatment
28	50	Diagnostic FL Diagnostic DLBCL Relapse DLBCL	Right cervical node Right cervical node Retroperitoneal node	Low grade N/A N/A	0 4(10) 6(9)	N/A	N/A

Table 3.7. *Additional clinical information regarding 6 patients who underwent IGHV targeted sequencing.*

3.4.9 N-gly motifs in the major clones of temporal biopsies

All patients had an *IGH-VH3* rearranged gene. With the exception of patient 27, patients contained at least one acquired N-gly motif site in their major clone which was retained throughout temporal samples (Table 3.8), consistent with the previous FL-tFL study described above.

Patient	Disease Event	VDJ rearrangement	Homology to V gene (%)	Region in IGHV	Motif aa sequence
23	Diagnosis	V3-30,D3-16, J6	89.24	CDR3	NFS
	Transformation		87.89	CDR3	***NVS***
24	1st relapse	V3-11,D3-16, , J1	85.65	CDR1, FR2, CDR2, FR3	NFS, NMS, NIT, NNS
	3rd relapse		86.1	CDR1, FR2, CDR2, FR3	NFS, NMS, NIT, NNS
	Transformation		85.2	CDR1, FR2, CDR2, FR3	NFS, NMS, NIT, ***NNT***
25	2nd relapse	V3-48, D1-26, J4	79.82	CDR2	NIS
	3rd relapse		79.82	CDR2	NIS
	Transformation		79.55	CDR2	NIS
26	1st relapse	V3-48, D3-10, J6	88.57	FR2, CDR3	NKS, NNS
	2nd relapse		88.57	FR2, CDR3	NKS, NNS
27	1st relapse	V3-23, D4-23, J6	89.11	CDR1	NFS
	3rd relapse	V3-23, D3-3*, J5*	83.47	FR3	***NLT***
28	FL diagnosis	V3-23, D5-18, J6	N/A	CDR3	NGS
	DLBCL diagnosis		N/A	CDR3	NGS
	DLBCL relapse		N/A	CDR3	NGS

Table 3.8. *N-gly motifs identified in the major clone of six FL patients taken at different time points of disease*. Italic red text refers to differences in the location or amino acid (aa) sequence of N-gly sites within the major clone across temporal samples. Major clones were determined through heteroduplex analysis following PCR amplification and verified from sequencing data, indicated by the highest count number. The multiple values in patient 24 and 26 relate to the multiple motifs observed. * Although temporal samples of patient 27 have different DJ rearrangements according to IMGT, when aligned they show highly similar CDR3 regions and share a t(14:18) breakpoint, indicating a clonal relationship.

Patient 23 samples were taken at diagnosis and transformation (1.5 years after diagnosis). In the major clone, one N-glycosylation motif was observed in the CDR3 region of the diagnostic sample, composed of the amino acid sequence NFS. The transformed major clone maintained the N-gly location in the CDR3 yet non-silent mutations within this region impacted the amino acid sequence (NFS>NVS). Samples for Patient 24 spanned the first relapse, second relapse and transformed disease events. In the major clone of the first relapse event, four N-gly motif sites were observed in several locations across the variable region, including CDR1 (NFS), FR2 (NMS), CDR2 (NIT) and

FR3 (NNS). All four sites were conserved in the 2nd relapse and transformed major clones, although the fourth site in the transformed sample had a different amino acid composition as a result of non-silent mutations (NNS>NNT). This conservation is despite the patient undergoing several lines of therapy in between biopsies (Table 3.7). Patient 25 samples, taken at 2nd relapse, 3rd relapse and transformation had a conserved motif site and composition within the CDR2 region. Interestingly, the CDR3 length in the transformed sample major clone has a nine nucleotide deletion in the N1 region resulting in a shortened CDR3 length of 14 compared to the 2nd and 3rd relapse clones which had a CDR3 length of 17 (Figure 3.12). Although CDR3 length conservation is normally indicative of clonality, we were able to establish a clonal relationship by events sharing identical BCL2 breakpoints and VDJ rearrangements (Table 3.8 and Figure 3.13).

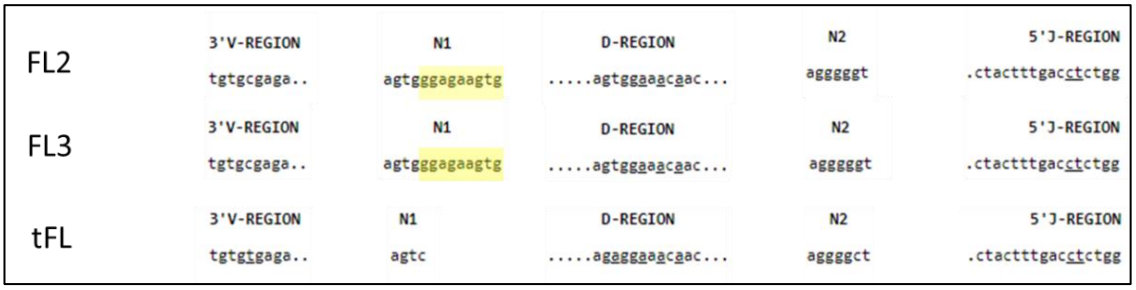


Figure 3.12: *Alignment of the major clone CDR3 region of Patient 25 disease events.* The nine nucleotides highlighted in yellow refer to the nucleotides which are absent in the N1 region of the transformed major clone, resulting in a shorter CDR3 length of 14 instead of 17 (FL2 and FL3), as indicated by IMGT-V-QUEST. The CDR3 length seen in the major clone is conserved in the significant majority of subclones within the specific disease event.

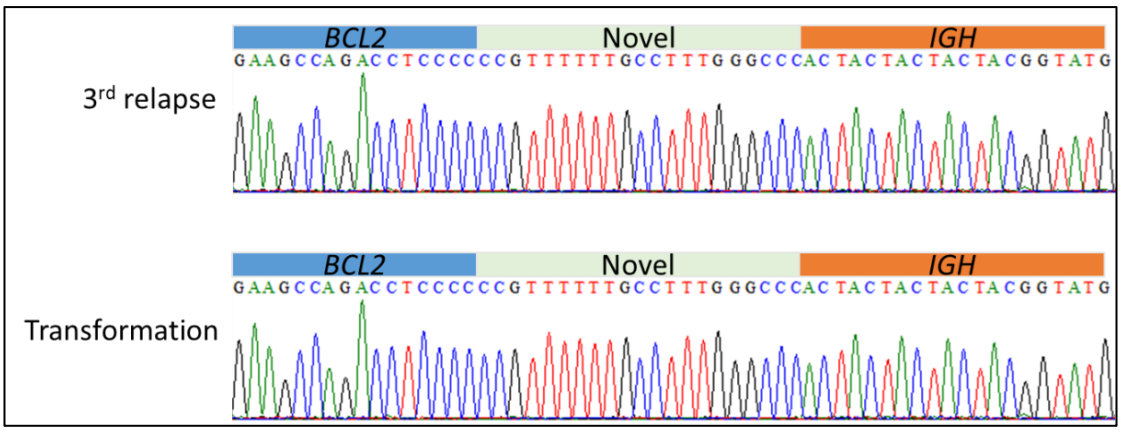


Figure 3.13: *BCL2 breakpoint of 3rd relapse and transformation events of patient 25.* Identical breakpoint nucleotide sequence indicates clonality.

For patient 26, the major clones of the 2nd and 3rd relapse samples contained two N-gly sites NKS and NNS located in FR2 and CDR3, respectively.

Patient 27 temporal samples were taken from different anatomical sites. 1st and 2nd relapse samples were taken from the right axillary node and right femoral node, respectively. Interestingly, the N-gly motif site observed in the 1st relapse major clone in the CDR1 region (NFS) was not found in the 2nd relapse major clone which contained a different motif site located in FR3 (NLT).

The first two events for patient 28 were taken from the same anatomical site, yet the relapsed DLBCL was taken from a distinct site. Unlike patient 27, the motif site and amino acid sequence (NGS) in the CDR3 region was conserved throughout progression.

Non-synonymous mutations lead to the acquisition of the asparagine residue in all patient samples, despite the presence of several germline encoded asparagines. The serine/threonine amino acid was either a feature of the germline or arose through synonymous or non-synonymous mutation (Table 3.9). The asparagine residue was generated through partial or complete codon substitution as highlighted in Patient 26, whereas across patients, the middle codon did not undergo variation except for in Patient 25. Samples with a C terminal serine were most commonly found in the germline and could be mutated to a threonine. Motifs located in the CDR3 region (as seen for patients 24, 28 and 26) were unable to be compared to a germline sequence, therefore

their origin from either synonymous or non-synonymous somatic mutations could not be established.

Patient	FL Disease Event	AA Change	Bp mutation[A]		
			1	2	3
24	1st relapse	TFS>NFS, YMS>NMS, YIS>NIT, KNS>NNS	AAC, AAT, AAC, AAC	TTC, ATG, ATT, AAC	AGT, AGT, ACT, TCT
	3rd relapse	TFS>NFS, YMS>NMS, YIS>NIT, KNS>NNS	AAC, AAT, AAC, AAC	TTC, ATG, ATT, AAC	AGT, AGT, ACT, TCT
	Transformation	TFS>NFS, YMS>NMS, YIS>NIT, KNS>NNT	AAC, AAT, AAC, AAC	TTC, ATG, ATT, AAC	AGT, AGT, ACT, ACT
25	2nd relapse	YIS>NIS	AAC	ATA	AGT
	3rd relapse	YIS>NIS	AAC	ATA	AGT
	Transformation	YIS>NIS	AAC	ATA	AGT
26	1st relapse	GKG>NKS	AAT	AAG	AGT
	2nd relapse	GKG>NKS	AAT	AAG	AGC
27	1st relapse	TFS>NFS	AAC	TTT	AGC
	3rd relapse	SLR>NLT	AAC	CTG	ACA

Table 3.9: *Analysis of somatic mutations giving rise to N-gly sites.* Red letters indicate a nucleotide replacement mutation.[A] refers to amino acid position 1, 2 and 3 of N-gly motif. Patients 1 and 6 were excluded from analysis due to motifs being located in the CDR3 region which does not have comparative germline sequence to determine somatic mutations at motif site.

3.4.10 Distribution of SHM within the subclonal repertoire

Figure 3.14 summarises the impact of somatic hypermutation across the variable region of Patients 23-25 by detailing the location and distribution of single nucleotide variants (SNV) in the subclonal repertoire, indicating the total numbers of non-synonymous versus synonymous mutations. Whilst there are shared distributions of specific transversion and transition mutations in certain regions across temporal samples, indicative of shared SHM in a common ancestral cell, there are also changes in distribution as the disease progresses as seen by the acquirement of new point mutations and absence of previous mutations. This is supportive of a divergent evolution between serial events.

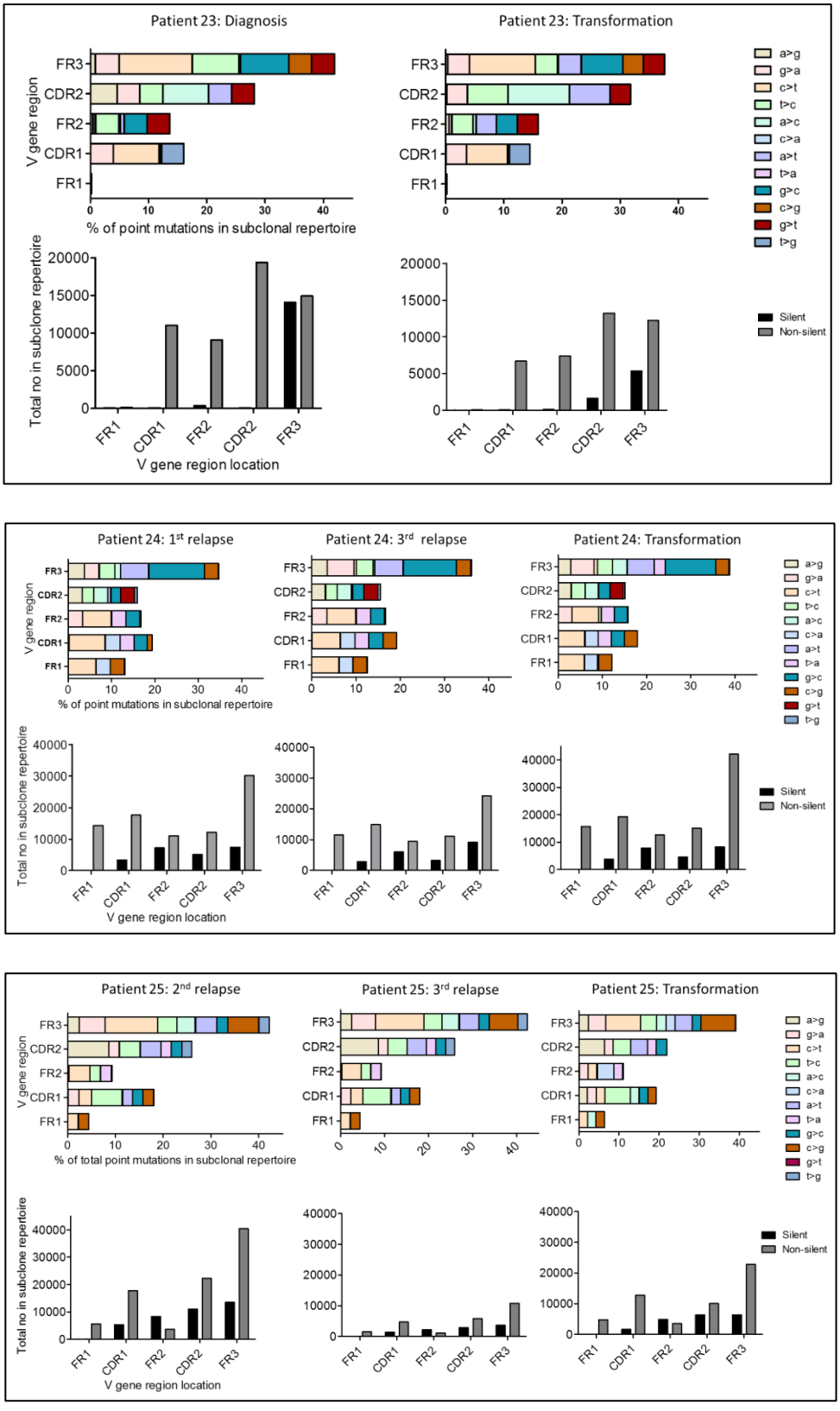


Figure 3.14: *Distribution of V gene nucleotide substitutions found in subclones across disease events for patients 23-25*. The CDR3 region has been excluded due to lack of a germline sequence. As the disease progresses, there are shifts in the mutation distribution.

Figure 3.14 also indicates how mutations are either silent or non-silent. BASELINe (Section 2.14) was used to determine whether the distribution of mutations in the immunoglobulin sequences the across the variable region was indicative to antigen-driven selection, as calculated by the ratio of replacement/silent mutations in the CDRs and framework regions (Figure 3.15). The 2nd relapse subclones of patient 25 showed positive selection of replacement mutations in their CDRs and negative selection of replacement mutations in their frameworks, which is typical of antigen-driven selection profile. This is because replacement mutations could potentially improve antigen affinity of the immunoglobulin and therefore are more likely to accumulate in the antigen binding associated CDRs. In contrast, replacement mutations in the framework regions could affect the structural integrity of the antigen-binding site of the immunoglobulin. In the subclone repertoire of successive disease events of patient 25 and in the temporal samples of all other patients, we did not observe positive selection in replacement mutations in the CDRs, as highlighted in Figure 3.15. Replacement mutations in the framework mutations were negatively selected in all temporal samples, which is expected as the surface immunoglobulin is constantly present and functional throughout FL disease. Comparatively, replacement mutations in the CDRs were more positively selected in all cases as indicated by negative $P_{CDR/FR}$ values. This finding highlights how subclones are selected based on structural integrity of the immunoglobulin but not by antigen, which is indicated in other studies. This infers an alternative selection pressure driving clonal evolution.

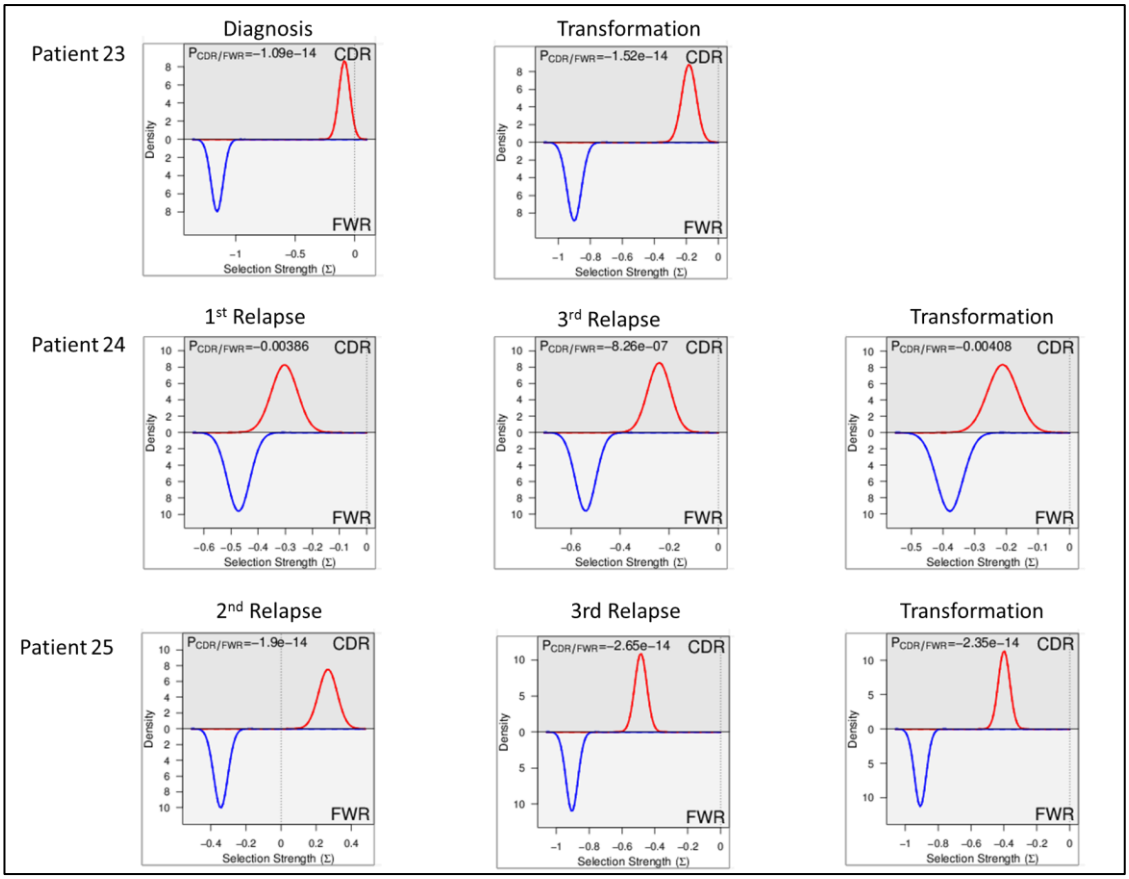


Figure 3.15: *Representative BASELINe graphical output of subclones from temporal samples of three patients*. BASELINe compares the frequency and location of replacement mutations against expected values from antigen-driven selection. Graphs highlight strong negative selection of replacement mutations in the framework regions but no positive selection in the CDRs (except for 2nd relapse subclones in patient 25, as indicated by a positive selection strength value). Negative $P_{CDR/FR}$ indicates CDR replacement mutations are more positively selected than framework mutations.

3.4.11 N-gly motif site conservation in subclone population

To elucidate whether N-gly acquirement is a clonal event, I interrogated the subclone population. For patients containing one motif site in their major clone, ≥97% of the subclone population within and across disease events maintained the motif sites (Figure 3.16). Strikingly, for Patients 24 and 26 which had multiple N-gly sites, no subclone with the complete absence of motifs was detected within and across disease events. For Patient 24, the 1st, 2nd and 3rd sites were conserved in >96% of subclones across events. The 4th site was conserved in 97.2% of clones in the 1st relapse sample and in 81% and 85% of subclones in the 3rd relapse and transformation samples, respectively.

Interestingly, for all patients no further motifs accumulated within or across event subclones that were not found in the major clone.

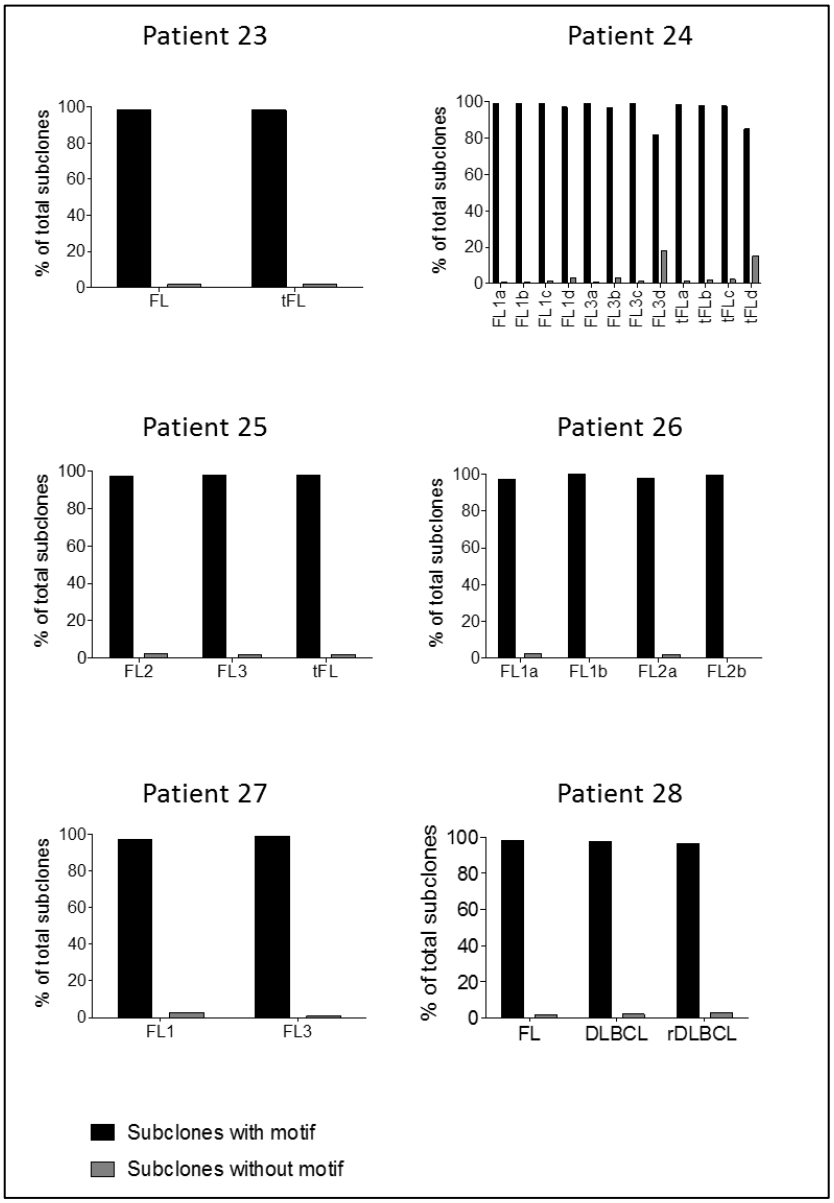


Figure 3.16: *Comparison between % of tumour subclones with presence and absence of N-gly motif site in the heavy chain variable region across disease events*. The motif location and sequence for each patient is detailed in Table 3.8. The majority of subclones across patients and disease events retain the N-gly site (*p*-0.0001). FL-diagnosis, FL1-1st relapse, FL2-2nd relapse, FL3-3rd relapse, tFL- transformation, DLBCL (diffuse large B cell lymphoma), rDLBCL (relapsed DLBCL). a-d in Patient 24 refer to the four N-gly sites located in CDR2, FR2, CDR3, FR3, respectively. a&b in Patient 26 refer to the two N-gly sites located in FR2 and CDR3.

3.4.12 Diversity within the N-gly motif site

The majority of subclones present the identical nine nucleotides that encode for the N-gly motif found in the major clone, indicating that the SHM that discriminates between these subclones lies within a different location of the variable region. However, for a relatively small proportion of subclones, there is diversity within nucleotide sequence encoding for the motif. This is expected as a consequence of the (largely) random process of SHM which does not differentiate between seemingly favourable and non-favourable mutations. Table 3.10 highlights the % of total unique subclones with a different codon sequence to that of the major clone. The actual number of affected clones is also indicated.

Patient	FL Disease Event	% of total subclones with different codon sequence in N-gly region	Clones with different codon sequence in N-gly region	
			Clones maintaining N-gly motif	Clones without N-gly motif
23	Diagnosis	3.92	61	46
	Transformation	6.51	79	31
24	1st relapse	2.72, 3.2,58.41, 0.06	61, 78, 2022, 128	35, 35, 40, 98
	3rd relapse	4.78, 7.88, 32.39, 24.3	124, 142, 944, 186	21, 97, 38, 551
	Transformation	3.38, 3.73, 9.33, 18.84	76, 75, 1279, 150	58, 73, 91, 597
25	2nd relapse	0.85	281	104
	3rd relapse	8.7	136	43
	Transformation	28.5	426	29
26	1st relapse	34.57, 0	26, 0	2, 0
	2nd relapse	4.84, 0.004	9, 0	5, 1
27	1st relapse	15.71	18	4
	3rd relapse	5.66	18	3
28	Diagnosis	36.1	906	40
	DLBCL diagnosis	5.64	194	71
	DLBCL relapse	1.81	86	128

Table 3.10. *Analysis of subclones not sharing the identical codon sequence of the N-gly motif found in the major clone.* % of the total tumour related subclones that do not contain the identical codon sequence seen in the major clone are highlighted on the right. The actual numbers of these subclones which do and do not encode for the N-gly motif are given on the right. The several values seen for Patient 24 and 26 are representative for the multiple N-gly sites seen for these patients.

Interestingly, an average of 77% of affected clones were positive for the motif site because of either synonymous mutations of the asparagine or

synonymous/nonsynonymous mutations of the serine or threonine residue. Only 23% of subclones were absent for the site, indicating a positive selection of subclones that maintain the motif site. This is indirectly supported by motif negative subclones only making up ≤1% of the total count number in samples expressing only one motif site (Supplementary Table 1.5).

The diversity within the nucleotide sequence encoding for the asparagine and serine or threonine residues in these positive subclones is shown for each disease event in Figure 3.17. There are twenty possible nucleotide combinations that can give rise to N-gly motifs when considering the first and last amino acids only. As seen across patients, a variety of distinct motif encoding sequences arises in subclones through SHM, with the most diverse range of sequences seen for the 3rd motif site of patient 24.

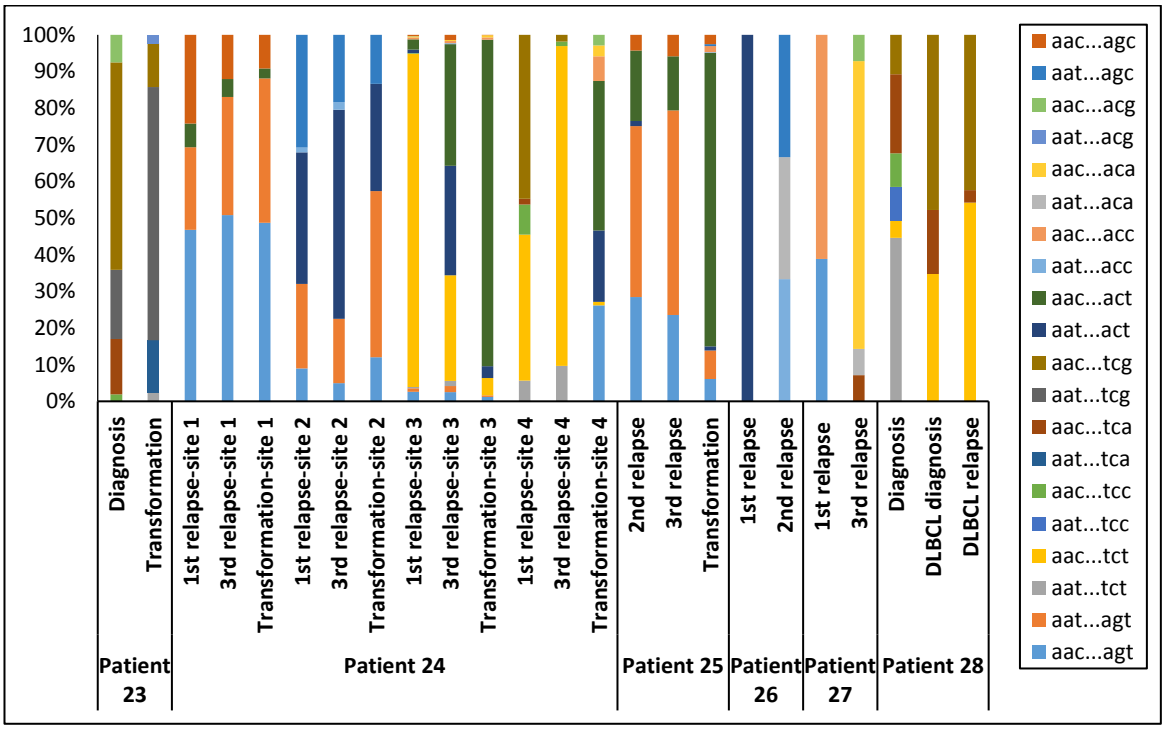


Figure 3.17: *Diversity in the nucleotide sequence of the first and last amino acid of motif positive subclones.* Subclones represented here did not display either the same asparagine or serine/threonine encoding nucleotide sequence as that of the major clone. The middle codon for all sequences was checked to ensure for the absence of proline as this middle amino acid would negatively affect the functionality of the N-gly motif.

3.4.13 Analysis of subclones with highest and lowest degree of SHM

The range of SHM is indicated in Table 3.11, in which the least and most mutated subclones for each disease event were identified by the IMGT High V-QUEST program. The % difference in homology between the least and most mutated subclones ranged from 0.8 to 13.3, depending on the patient and disease event. In some instances, the degree of SHM (indicated by these extreme subclones) does not increase as the disease progresses. This is reflected in patients 23-25, in which the least mutated transformed subclone has a greater semblance to the germline when compared to the least mutated subclone of the preceding FL events. The subclones highlight that evolution is not a linear process with the accumulation of SHM over time and supports the divergent pattern seen in other FL SHM studies (Carlotti et al., 2009, Carlotti et al., 2015).

When we examine these extreme subclones (least and most mutated) collectively for each patient, we gain an insight into the point at which N-gly motifs were introduced. For example, in the transformed event for patient 24, we observe that the fourth motif site has been acquired in a subclone which has at least 93.2% homology to its germline, translating into just 17 nucleotide substitutions. However, further SHM is necessary for the acquirement of sites 1-3. For patients 23 and 26, motifs are acquired within their least mutated subclones, however, for patient 25 this isn't the case. It suggests that the degree of SHM required for motif acquirement varies between patients and could be influenced by the *V* gene used. However, once acquired, single motifs are rarely lost despite a high level of ongoing SHM, as evidence in the most mutated subclones of patients 23, 25, 26 and 27. Although patient 24 has the loss of the 1st and 2nd motif sites as a result of ongoing SHM, they retain the 3rd and 4th site in their most mutated subclones.

	Patient 23		Patient 24			Patient 25			Patient 26		Patient 27		Patient 28		
Least mutated subclone	FL	tFL	FL	FL2	tFL	FL2	FL3	tFL	FL1	FL2	FL1	FL3	FL	tFL	tFL1
% homology to G.L *V* gene sequence	89.23	89.6	87.6	87.1	93.2	78	79.8	89.2	89.4	88.6	98	89.1	97.8	93	92.3
N-gly motif presence (Y/N)	Y	Y	Y, Y, Y, Y	Y, Y, Y, N	N, N, N, Y	Y	Y	N	Y,Y	Y, Y	N	Y (FL1 motif)	Y	Y	Y
Most mutated subclone															
% homology to G.L *V* gene sequence	80.2	80.7	79.9	80.7	79.9	74.4	74	79.1	88.6	86.9	87.5	81.5	79.4	71.3	73.5
N-gly motif presence (Y/N)	Y	Y	N, N, Y, Y	N, N, Y, Y	N, N, Y, Y	Y	Y	Y	Y,Y	Y,Y	Y	Y	Y	Y	Y
% difference in homology between least and most mutated subclones	9.03	8.9	7.7	6.4	13.3	3.6	5.8	10.1	0.8	1.7	10.5	7.6	18.4	21.7	18.8

Table 3.11: *N-gly motif status in most and least diverse subclone based on degree of SHM compared to V gene germline sequence.* Sequences were analysed for presence or absence of the N-gly motif site (s) found in the major clone.

3.4.14 Discordant SHM in distinct anatomical sites

When comparing the subclones of patient 27, taken from two lymph nodes at distinct anatomical sites, we observe a clear discordance in the SHM pattern of the two populations and how this translates into a highly distinct amino acid sequence in the variable region of the major clones (Figure 3.18). This suggests an early divergence in the evolution of subclones giving rise to these two events. The presence of distinct motifs in each event (Table 3.8) infers that acquirement of motifs occurred after divergency in which event specific subclones underwent unique SHM processes. Despite this difference and the selection pressures imposed by microenvironmental difference, tumour subclones are universally reliant on N-gly motifs.

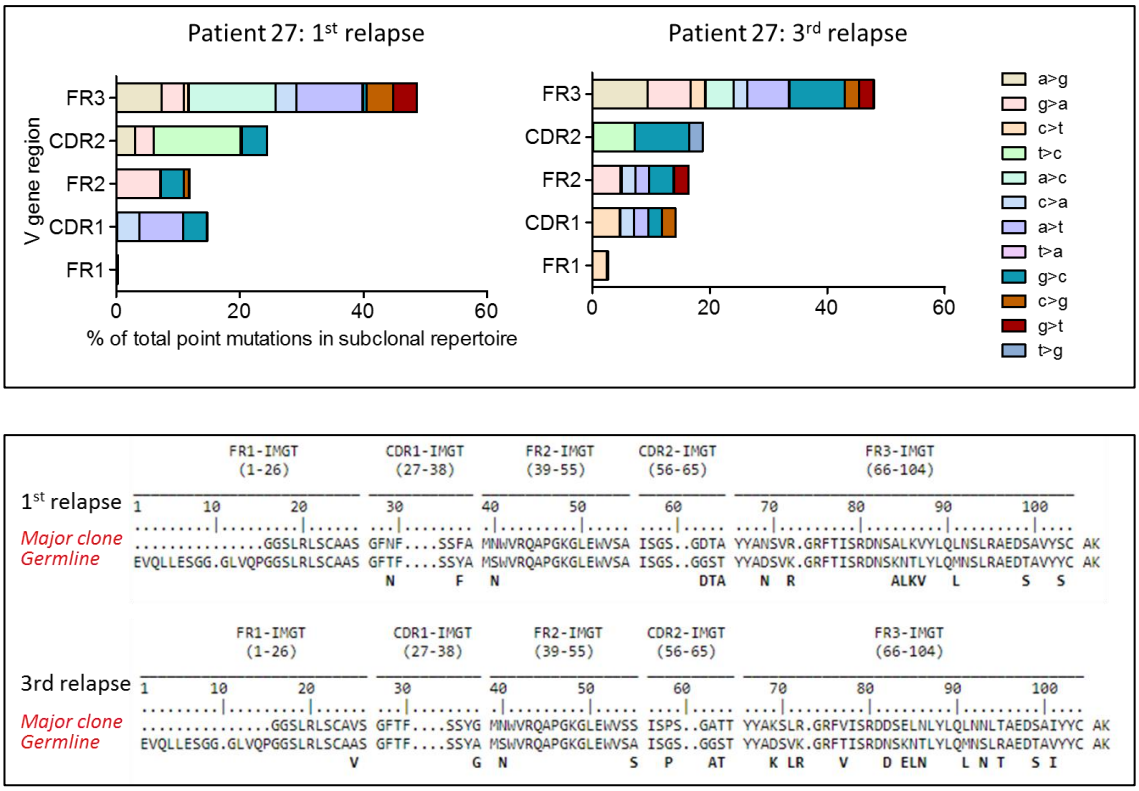


Figure 3.18: *Discordance in the SHM pattern of subclones identified in temporal samples from patient 27 across the variable region.* The *V gene amino acid sequence of the major clones for the serial samples is given below.* Major clone and germline amino acid sequences are indicated, with amino acid changes from the germline sequence highlighted in bold letters below. The discordant amino acid changes reflect the outcome of the distinct SHM pattern in the two events. Yet a clonal relationship was previously established between the two, indicating an early divergence in evolution (Carlotti et al., 2015).

3.4.15 Fate of motif negative subclones during progression

The subclone population changes as the disease progresses, influenced by pressures that select for the outgrowth of subclones containing advantageous mutations that confer a survival advantage. These subclones may be present in the preceding event maintaining a background 'role' before acquirement of further genetic aberrations causes its proliferation and dominance. Therefore, the analysis of shared subclones across temporal samples can give us insight into the subclones important in driving disease progression. Subclone populations were merged across disease events to identify candidate clones (Figure 3.19). With the exception of Patient 25, a small number of subclones are shared across events. These subclones are long-lived, as exampled in Patient 27 in which the two biopsies were taken 6.75 years apart (Table 3.7). It was interesting to observe that subclones were shared between distinct anatomical sites as seen in patient 27 and in patient 28, in which the relapsed DLBCL was biopsied from the retroperitoneal node in contrast to the first two events biopsied from the right cervical node.

Analysis of the shared subclones revealed they were all motif positive, including those trafficked between anatomical sites. For patient 27, the subclones contained the motif of the second event. This indicates a trafficking ability of motif positive subclones. Patient 25 was an exception as 1.5% of the shared clones were negative (n=20). Interestingly, most of these negative clones have a higher relative count number in the successive disease event, suggesting that they confer an advantage (Supplementary Table 1.6). The lack of shared motif negative subclones in all other patients indicates that progression subclones are dependent on N-gly motifs.

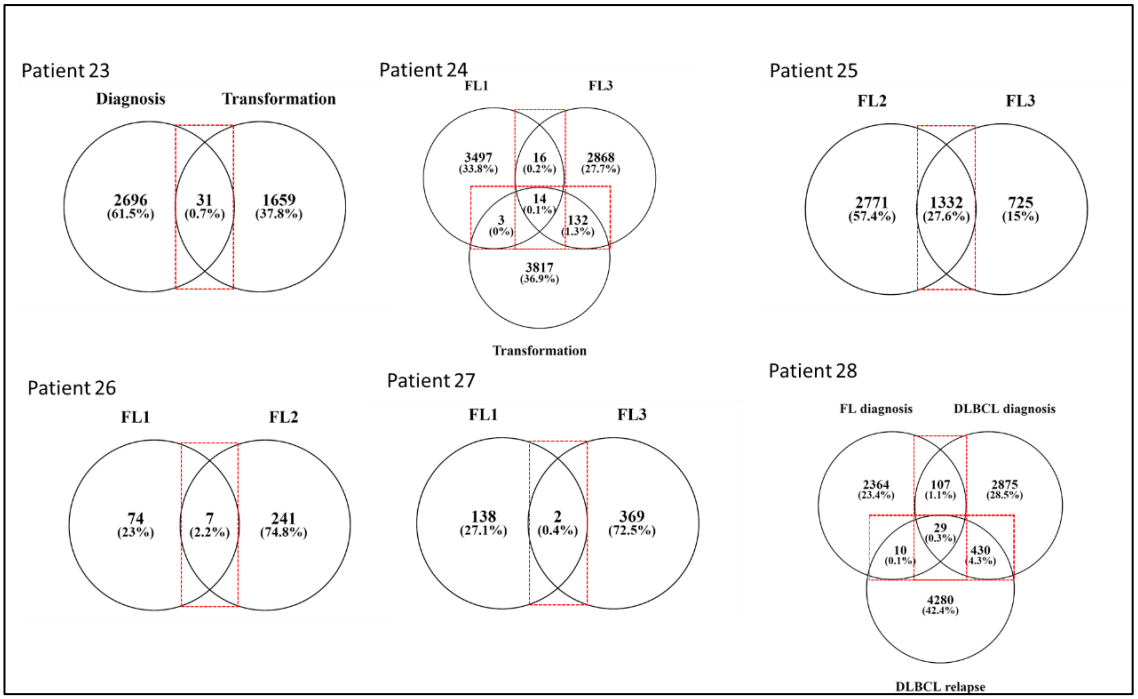


Figure 3.19: *Venn diagrams showing the number of shared and distinct subclones across disease events.* Numbers inside circles reflect a number of subclones for the particular event with % of the total subclone repertoire across events indicated in brackets. For Patient 25, the transformation sample was omitted from analysis due to the difference in CDR3 length in comparison to the 1st and 2nd relapse.

3.4.16 Lineage trees reveal origin and fate of N-gly negative subclones

To investigate whether N-gly motifs are acquired early on in the disease, we investigated the evolution of subclones based on their *IGHV* sequences. The IgTree program is specifically designed for variable region immunoglobulin sequences and calculates the minimal and most probable mutational pathways separating subclones, highlighting intraclonal hierarchy (Barak et al., 2008) (Section 2.24). For Patients 26 and 27, the complete hierarchy can be visualised in lineage trees (Figures 3.20 and 3.21). Due to a the large number of subclones detected for Patients 23 and 25, only the negative subclones, their parent clone and their progenitor clones are depicted in Figure 3.22 to highlight their generation and contribution to the clonal repertoire within an event. As Patient 24 did not have any truly negative subclones, the analysis was omitted.

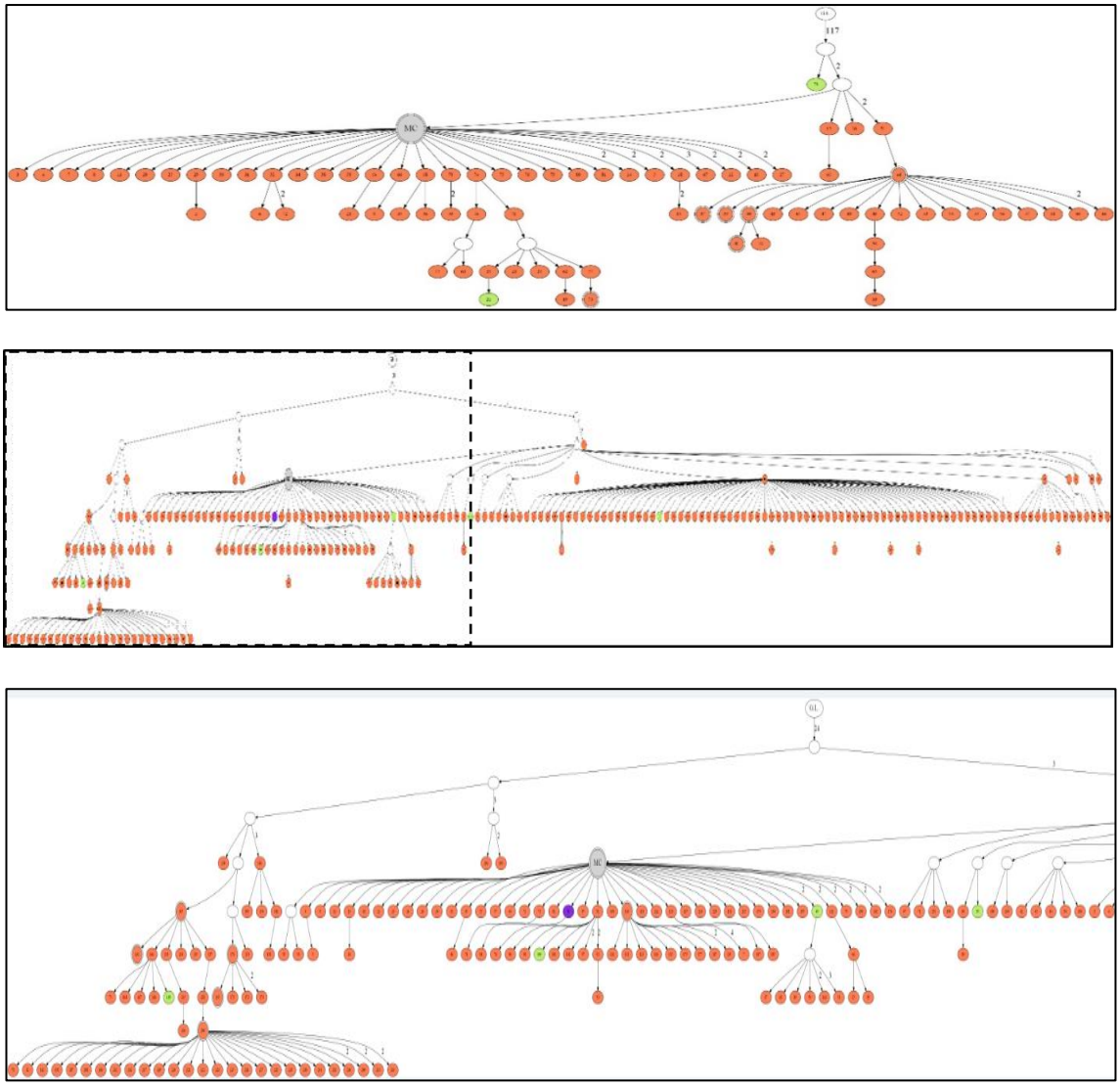


Figure 3.20: *Lineage tree for the two disease events of Patient 26.* The top panel represents the FL disease whereas the middle panel represents the second relapse, FL2. The bottom panel represents the zoomed in areas of FL2 (indicated in dotted box within second panel) Orange circles:clones with both motif sites; purple circles:clones with only the first motif site; green circles: clones with only the second motif site; white circles: clones inferred to exist but not detected through 454 sequencing; grey circle: major clone containing both sites. Clones that have lost a motif site in either the FR2 or CDR3 regions are able to regain the site following ongoing SHM (boxes in bottom panel). No new motif sites are acquired in clones.

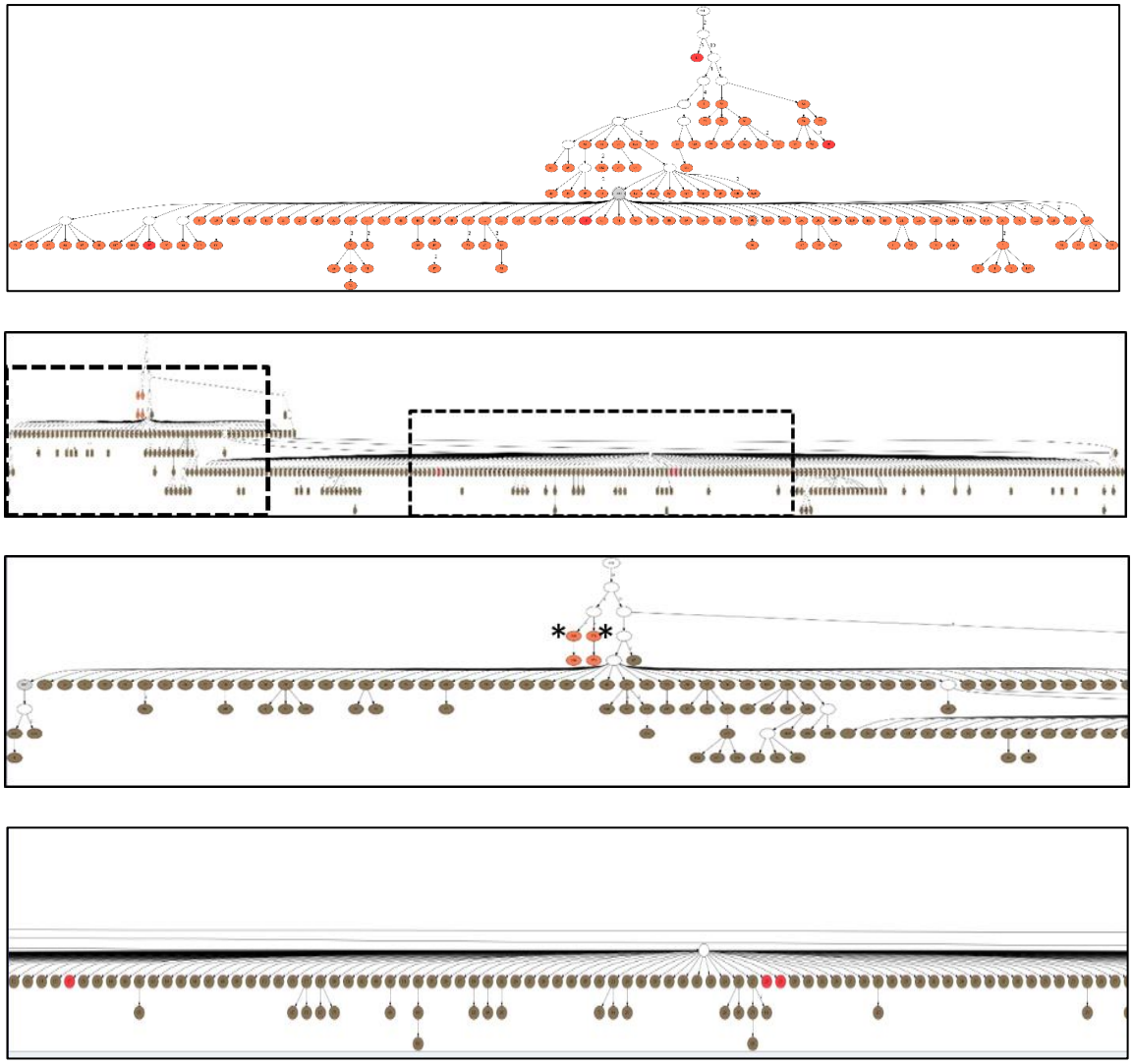


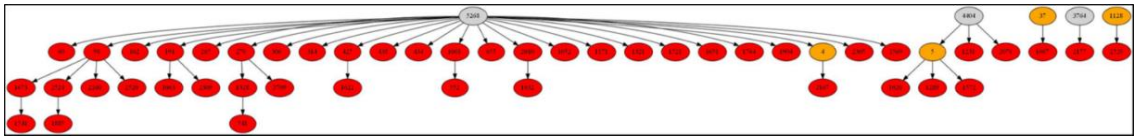
Figure 3.21: *Lineage tree for the two disease events of Patient 27.* Top panel: FL disease event. Second panel: Second relapse event, FL2. Third and fourth panel: Zoomed in areas of FL2 (indicated in boxes within second panel) Orange circles: clones with motif site found in FL event; brown circles: clones with motif site found in FL2 event; red circles; clones without a motif site. Clones without motif do not give rise to further subclones and cannot regain motif sites unlike in multiple N-gly site tumour subclone populations (Patient 26). *Represents subclones containing the motif site of the preceding disease event.

For patient 26, the earliest experimentally derived subclones (identified as filled circles closest to the germline Ig sequence at the top of the tree) are motif positive. While patient 26 does not have any truly negative subclones, several subclones lose at least one motif site. These subclones are observed to undergo further SHM and re-acquire the lost site (Figure 3.20), giving rise to several further clones. This is in contrast to patient 27 in which the loss of the single motif results in the subclone not undergoing

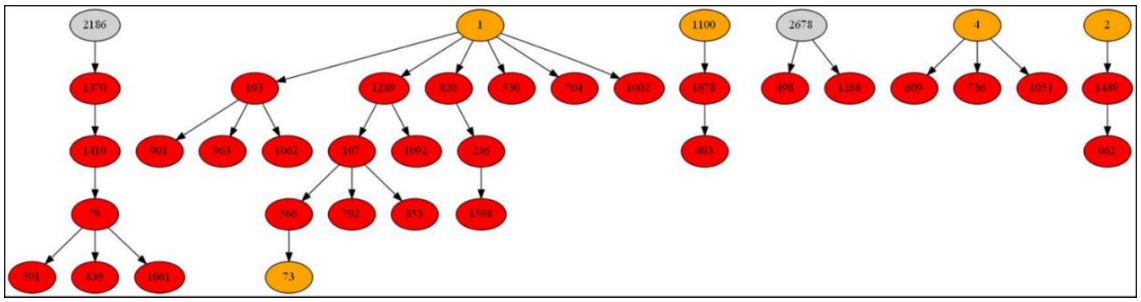
further diversification or expansion (Figure 3.21). As these clones are lost from progression samples, we can infer their elimination. All motif negative subclones in patient 27 were descendants of motif positive clones as a result of ongoing SHM and were not early events as suggested from their downstream positioning within the lineage trees.

With greater numbers of motif negative subclones owing to an increase in overall subclones, Patients 23 and 25 lineage trees give a more comprehensive insight into the behaviour of negative clones in the tumour hierarchy (Figure 3.22). However, it is important to note that these negative subclones represent a minority within the heterogeneous population. For patient 23, negative clones represented 1.7% and 1.8% of the subclone population in diagnosis and transformation events, respectively. For patient 25, negative clones found in 2nd relapse, 3rd relapse and transformation represented 2.5%, 2.1% and 1.8% of the population, respectively. Motif negative clones were found to arise from either a positive or negative clone, through a single nucleotide variant. Several negative clones can arise from a shared positive ancestor, as depicted through the wide branching. Except for one clone in Patient 23, further SHM in these negative clones does not result in site re-acquirement or gain of new sites. However, these clones make up on average 0.01% of the total count in all patient samples.

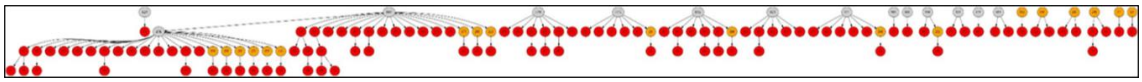
Patient 23: Diagnosis



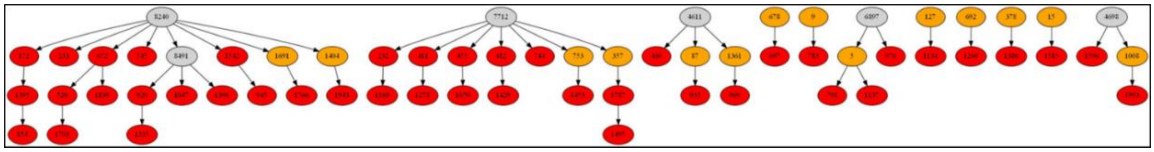
Patient 23: Transformation



Patient 25: 2nd relapse



Patient 25: 3rd relapse



Patient 25: Transformation

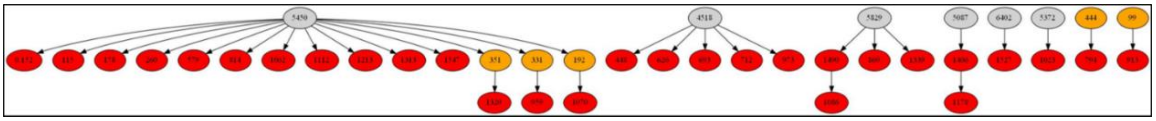


Figure 3.22: *Hierarchy of motif negative clones in relation to their direct ancestral clone for Patient 23 and 25.* This diagram represents part of the lineage trees as the majority of subclones (which are positive) are not shown and the interconnection between clones from a germline sequence is not depicted. Red clones represent motif negative clones. Orange clones represent motif positive clones derived from the sequencing platform. Clones in grey represent clones not detected by the sequencing platform but predicted to exist by the IgTree program. They are all assumed to contain the motif site due to a number of other experimentally detected progeny clones being motif positive. Numbers inside clones are generated through the IgTree program to numerate clones.

3.4.17 N-gly motifs in evolution of composite Hodgkin's Lymphoma and FL

We next wanted to investigate whether N-gly motifs were an early acquirement in other ancestral clones giving rise to more phenotypically and pathologically distinct disease events. Three cases of composite Hodgkin's lymphoma (HL) and FL were analysed (Figure 3.23) which had previously been identified as originating from a common cell of origin through shared VDJ rearrangements and SHM patterns (Marafioti et al., 1999, Brauninger et al., 1999, Kuppers et al., 2001). For case 1, a 38 year old male was originally diagnosed with HL in the cervical lymph node. Following treatment, the patient developed FL in the same anatomical site two years later. For case 2, a biopsy of retroperitoneal nodes revealed composite FL and mixed cellularity HL with abundant Reed Sternberg cells (RSCs) in a 75 year old female. For case 3, a male was diagnosed with composite FL and HL in the splenic hilar lymph node.

IGHV sequences for the HL cell and least mutated FL cell were obtained from the ENA website, as described in section 2.21. The least mutated FL sequences were used to minimise the evolutionary distance to the common cell of origin. Unlike FL, HL cells do not have intraclonal heterogeneity regarding their *IGHV* sequences and therefore the HL sequence used represented all HL tumour cells.

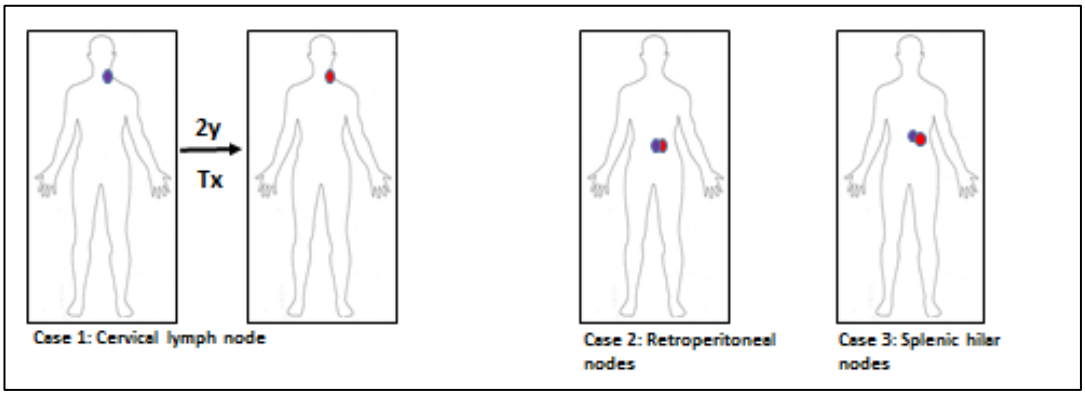


Figure 3.23: *Patients used in composite HL and FL IGHV analysis*.

For case 1, the FL and HL *IGHV* had one shared nucleotide substitutions, with the RSC and FL cell going on to gain an additional one and fifteen mutation(s), respectively

(Figure 3.24), indicating an early divergency between the two events. Through unshared SHM, the FL clone acquired two N-gly motifs NYS and NRS in the FR3 and CDR3 regions, respectively. However, RSC did not contain acquired motifs. Cases 2 and 3 had a greater number of shared SHMs compared to case 1, with 55 and 28 substitutions, respectively. In both cases, both the RSC and FL cells contained an identical N-gly motif, indicating acquirement occurred through shared SHM before divergency occurred to give rise to the distinct disease events. For case 2, the motif NMS was located in the CDR2 region and for case 3, the motif NIS was also found in the CDR2 region.

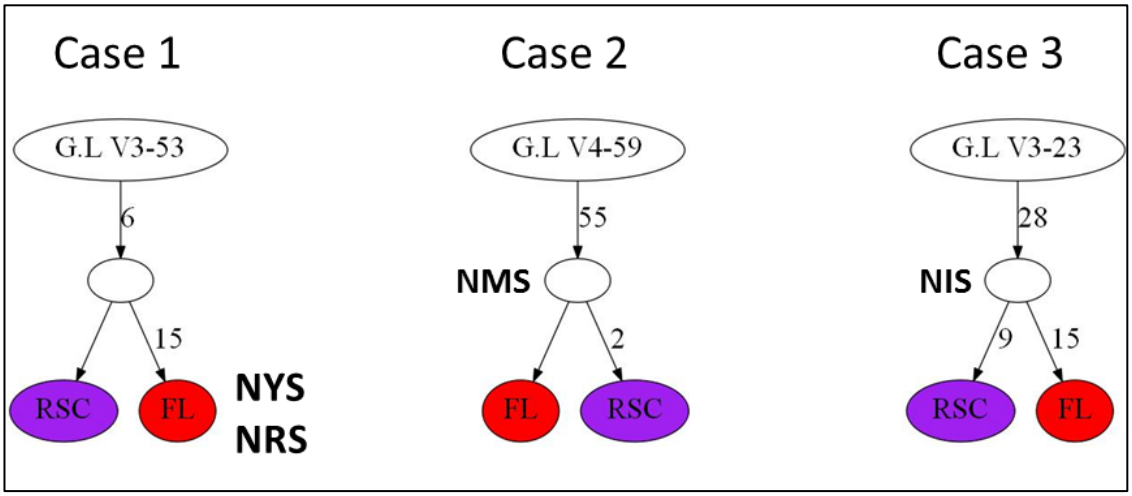


Figure 3.24: *Clonal relationship between composite HL and FL based on IGHV sequences*. For case 1, divergency between clones specific for disease events occurred early, as indicated by the low number of shared SHM. For case 2 and 3, divergency occurred at a later time point as indicated by high numbers of shared SHM. All three FL cells contained an N-gly motif, however motif presence in the RSC was dependent on the time point of divergency.

Although the distinct lymphomas are derived from the same germinal centre B cell, motifs are conserved only in patterns of late divergency, based on the relatively high number of shared *IGHV* mutations (Figure 3.25).

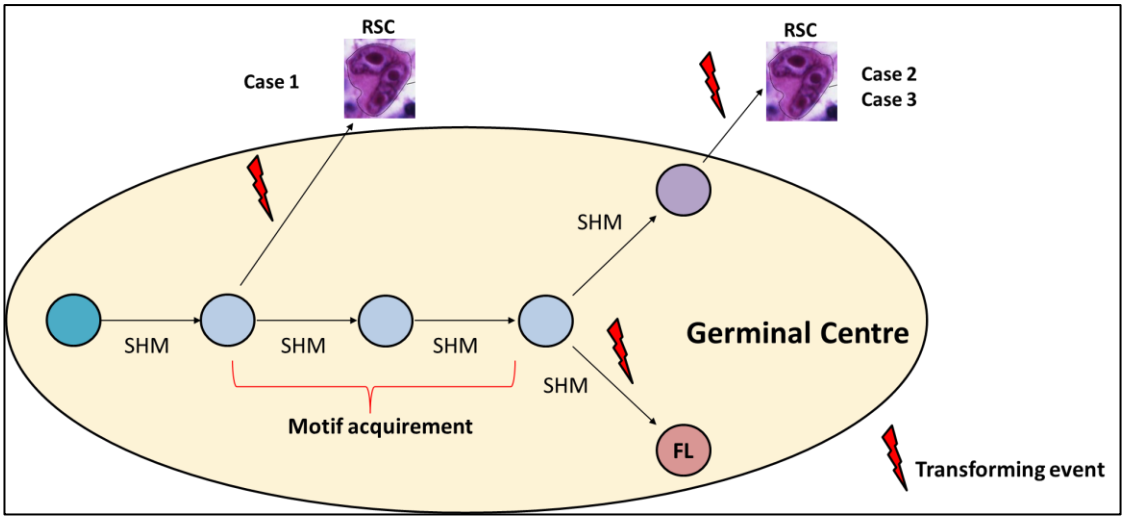


Figure 3.25: *Evolution model of composite FL and HL cases based on IGHV sequences.*

3.5 Discussion

FL patients undergo multiple cycles of relapse, eventually becoming refractory to therapy. This suggests that current therapies are not successfully targeting the early aberrations needed to propagate disease and as a result, the tumour can repopulate, gaining additional mutations that increase its complexity and reduce effective treatment options. Therefore, uncovering features of these ancestral cells and their microenvironmental interactions may offer durable outcomes for patients.

96.7% of the diagnostic cohort contained at least one N-gly motif within the *IGHV*, supporting the high prevalence of motifs seen in disease. The remaining 3.3% were negative for motifs. Comparison to the germline sequence of closest homology revealed that motifs were acquired. Whilst CDR3 located motifs do not have a comparative germline sequence owing to the random nature of somatic recombination, the lack of CDR3 motifs in normal somatically mutated B cells is strong evidence of their acquirement through SHM. We observed bias in the *V* genes used across patients, with *V3-48*, *V3-11* and *V3-23* being the most recurrently used amongst patients. This has also been observed in other studies looking at *V* gene distribution amongst FL samples (Berget et al., 2015, Aarts et al., 2000). Interestingly these *V* genes contained 'starter'

sequences that in the majority of cases, gave rise to N-gly motif sites through one or two nucleotide substitutions, indicating a reason for such *V* gene bias in disease.

Analysis of I*GHV* in paired FL/tFL samples revealed that motifs are conserved in the major clones of disease progression without additional sites being gained in transformed samples. Clonality of temporal events was confirmed by serial samples sharing identical V,D,J segments and t14;18 translocation breakpoints. For patients 12 and 22, the N-gly motifs were conserved in transformed samples despite unshared SHM occurring within the motif region as demonstrated by the change of amino acid sequences from NIT and NVS to NKS and NFS, respectively. The selection of a major clone which conserves the motif site indicates that motif positive clones have an advantage during disease progression. Conservation of motifs was despite patients undergoing therapy between biopsies, as highlighted in patient 15. We would expect therapy to exert a selection pressure on subclone in which those with advantageous mutations go on to dominate the clonal landscape. However, despite the assumed changes in the mutational profile of the major clones, conservation of sites is indicative of an ongoing reliance of the major clone on N-gly motifs.

A few germline *V* genes have naturally occurring N-gly motifs, including *V1-08*, *V4-34 and V5a* (Zhu et al., 2002). *V4-34* has the motif NHS, which is found in the CDR2 region. Therefore, it is intriguing that we do not observe a bias towards these specific *V* genes in disease. In fact, the natural motif site of the *V4-34* gene of patient 11 was found to be lost through shared SHM within our paired samples and a new site gained by shared SHM in the CDR3 region. This has been shown in other studies looking into SHM of FL cells (Zhu et al., 2002, McCann et al., 2008). Zhu *et al* observed six out of seven FL patients containing the *V4-34* gene lost the natural site with acquisition of a new site in either the CDR1 or CDR2 regions. For the one patient who retained the site, a new site was acquired in the CDR1 region. McCann *et al* found that the *V4-34* natural site was also lost in molecular clones derived from FL patients. Analysis of the variable region domain revealed that while acquired motifs were oligomannosylated, the natural site had no added sugars, suggesting differential glycosylation profiles of natural versus acquired motifs. This may explain as to why the natural site was lost in patient 11. Without mannoses binding, the FL cell cannot form interactions with lectins and

therefore the natural motif serves no functional relevance to the tumour cell and is eliminated.

FL is a highly heterogeneous disease with thousands of subclones present based on either distinct *IGHV* sequence or distinct mutational profiles. A minor subclone can gain dominance in a successive disease event if it possesses or acquires an advantageous mutation and therefore analysis of subclones is essential with regards to the study of disease evolution. Sequencing of up to ten subclones derived from cloning enabled us to analyse N-gly motifs within intraclonal variants. We found motifs observed in the major clone were also conserved in the subclones. This alluded to the idea of motif acquirement being a clonal and early event in disease evolution. However, to better address this required a comprehensive and ultra-deep analysis of the heterogeneous clonal landscape of disease which could only be assessed through next-generation sequencing.

We report for the first time and in unparalleled depth, the behaviour of N-gly motifs during disease progression by analysing the clonal repertoire of temporal t(14:18) positive FL samples based on amplicon sequencing of the *IGHV* gene. Samples ranged from diagnosis to transformation and included a mixed patient cohort with variable clinical disease courses and therapeutic interventions, reflecting the heterogeneous nature of the disease. In support of the high prevalence of N-gly motifs in FL, all patients harboured at least one acquired site in their earliest disease event major clone. We found that these sites were a clonal feature as seen through significant conservation both in the heterogeneous subclonal population and the overall tumour mass.

Sites are also retained in progression samples. These are seen in both progression exclusive subclones and shared subclones and although the acquirement of additional 'driver' mutations through natural or therapy-related selection pressures may dampen the tumour's microenvironment dependency, the motifs are suggested to retain an important functional significance at later stages of the disease, presumably mediated through the mannose-lectin interaction**.** This conservation occurred in both 'rich' and 'sparse' temporal samples, suggesting that the motifs are acquired at an early disease time point, before genetic divergency occurred to give rise to distinct genetic subclones.

Not all sites are simply maintained during progression. Some, like in patient 24 and 26 undergo SHM, resulting in changes to the nucleotide sequence (Table 3.9) which still encode for an N-gly site. This is similar to what we observed in patients 12 and 22 temporal samples.

The BASELINe program revealed that there was no positive selection in replacement mutations in the CDRs of any of the patient samples, except for the 2nd relapse sample of patient 25. For immunoglobulins that have undergone antigen-driven selection, replacement mutations in the CDRs could potentially improve antigen affinity, which is evidenced by the accumulation of mutations within these regions. However, all clonal populations did have strong negative selection against replacement mutations in the framework regions, providing strong evidence that subclones are selected based on structural integrity of the BCR but not by antigen recognition. This finding is supported by Hershberg *et al* and Zuckerman *et al* who also found strong negative selection against replacement mutations in the framework regions but no positive selection in the CDRs of FL variable regions (Hershberg et al., 2008, Zuckerman et al., 2010), inferring an alternative selection pressure driving clonal evolution.

To ensure that N-gly motifs were not conserved due to a lack of further SHM within the site, we analysed diversity within the nine base pair region. Whilst the majority of subclones shared the identical nucleotides to that of the major clone, a minority of subclones had a different sequence, which is expected due to the (largely) random nature of SHM which does not discriminate been favourable and non-favourable mutations. Strikingly, an average of 77% of affected subclones maintained the motif site through multiple codon combinations, indicating a positive selection.

Analysis of the least and most mutated subclones based on their homology to the germline *V* gene sequence gave us insight into the point at which N-gly motifs were acquired. In some cases, we were able to observe the motif within the earliest (least mutated) identified subclone whereas in others, this was not the case. This infers that the degree of SHM required for motif acquirement varies between patients and could be influenced by the *V* gene used and presence of so-called 'starter' sequences. However, once acquired, single motifs are rarely lost despite a high level of ongoing SHM

as evidenced in the most mutated subclones. Again this points to a selective advantage of tumour cells harbouring sites.

Negative subclones made up only a minor population across patients and were not passed on to successive disease events, in contrast to motif positive subclones. Lineage trees enabled us to reveal how this rare population manifests. All negative clones are derived from a positive clone, suggesting acquirement of motifs is an early event. The presence of negative clones is an expected occurrence as SHM however they represent only a small % of the tumour population and bulk, and are likely to be outcompeted by motif positive clones, perhaps due to loss of the microenvironmental interaction provided via the added mannoses. The compact branching between negative clones suggests that even in the absence of motifs, clones can still undergo SHM but do not reacquire sites or gain new ones in their progeny. However, it's important to take note that SHM does not indicate longevity of these subclones; they are separated by a single SHM event, suggesting they can arise within a short time frame of each other (e.g. hours or days). Indeed their absence from subsequent samples suggests they are short-lived, having adequate time for a single SHM event to give rise to a progeny subclone.

Patient 27 provided an interesting case for two reasons; the different anatomical sites for the two events and the discordant SHM within the *IGHV* between the two clonal populations. This discordance suggests an early divergence, where an ancestral cell with limited SHM, migrated from one site to another where selection pressures drove the outward growth of subclones with a specific SHM pattern. However, despite *IGHV* sequence heterogeneity, the acquirement of motifs within each population at different locations illustrates that motifs are an essential feature of FL, regardless of microenvironment differences. The sharing of two subclones with the motif of the diagnostic sample (Figure 3.19) highlights the trafficking ability of motif positive subclones, not just across disease episodes but between anatomical sites. The trafficking of only motif positive subclones between sites suggests a dynamic mannose-lectin interaction enabling GC exit and migration. This exclusive trafficking of motif positive subclones was also seen in patient 28 in which sampling of the diagnostic FL and sequential relapsed DLBCL occurred at distinct anatomical sites. However, for patient 28, SHM patterns between events were highly similar and the CDR3 N-gly site was

conserved throughout suggesting a late divergency between events from a shared ancestral cell**.**

Finally, we explored whether N-gly motifs were acquired within the common ancestral cell of disease events which are phenotypically distinct. Composite lymphomas are defined by either the simultaneous or sequential presence of two distinct lymphomas in a patient (Jaffe et al., 1994). Through their analysis, we can discriminate the events leading to lymphomagenesis in general and the events giving rise to each particular disease. We analysed the *IGHV* sequences of three cases of composite FL and HL. HL is characterised by occurrence of rare RSCs embedded in a mixed cellular infiltrate (Farrell and Jarrett, 2011). RSCs have a very different profile to FL cells including a large cell size, frequent multinuclearity, prominent eosinophilic nucleoli, loss of surface immunoglobulin expression and loss of B cell markers CD10, BCL-6 and CD20 (Kuppers et al., 2003). We found that in two of our cases, identical N-gly motifs were found in the RSC and FL cells, arising as an event of shared SHM and therefore a feature of a shared common ancestral cell. For case 1, only the FL cell acquired N-gly motifs, indicating N-gly motifs were not a shared event of the common ancestral cell. This is likely due to transforming genetic events leading to HL manifestation occurring early in evolution as indicated by the relatively low level of SHM seen within the RSC. Due to the lack of surface immunoglobulin expression on RSCs, it is unlikely that the N-gly motifs seen in case 2 and 3 have a functional and beneficial role within these cells. Instead, they are more likely to be an 'imprint' of a long duration of the common ancestral cell within the GC before transforming mutations were acquired.

The conservation of motifs in and across FL disease events and the lack of accumulation during ongoing SHM suggests they are an early and stable event in FL pathogenesis. Early events are usually determined through their conservation within temporal samples and for patients 23-25, WGS/WES had previously identified key genetic aberrations within a putative ancestral cell, known as the common progenitor cell (CPC). For all three patients this included mutations in *CREBBP*, *KMT2D* and *TNFRSF14*. Patients 23 and 25 had a 'sparse' CPC due to the lack of shared genetic aberrations across temporal samples, suggesting an early divergency with episodes arising from more genetically independent pathways. However, despite this mutational heterogeneity between

events, motif sites are conserved, identifying an important and rare feature of the CPC. This is an important finding as the CPC is believed to be the reservoir pool from which successive disease events arise, accounting for the high relapse rates experienced by the majority of patients (Okosun et al., 2014). The longest interval between the first and last events is seen in patient 28 with an interval of 6.75 years, suggesting an N-gly motif encoding CPC able to remain dormant for many years before a mutational event leads to a selective advantage and emergence of a new disease episode. The mannose-lectin interaction may enable tumour retention and survival of the CPC permitting the accumulation of genetic events that lead to overt disease, suggesting a critical priming event in FL manifestation. Although epigenetic deregulation is a considered CPC event as evidenced in the previous genetic profiling of Patient's 23-25 samples, our data imply that it is not solely sufficient for 'driving' the disease. Instead, the N-gly motif profile determines which clones are able to expand and survive during disease progression, irrespective of the genetic profile of the subclones which are believed to harbour such 'driver' mutations. Analysis of the genetic profile of motif negative subclones will determine the validity of this hypothesis.

Although the t14;18 translocation can be found in healthy B cells which do not go on to become malignant, N-gly sites in the variable region are restricted to GC derived lymphomas indicating an attractive and tumour specific therapeutic target which may lead to the loss of a critical CPC-microenvironmental interaction and reduce the frequency of relapse. The presence of N-gly motifs in the presumed FL precursor lesion, ISFN, supports the theory of N-gly motifs occurring at an early stage of pathogenesis (Kosmidis et al., 2017, Mamessier et al., 2015), being acquired even before disease manifestation. Figure 3.26 summarises how N-gly motifs impact the evolution of the disease.

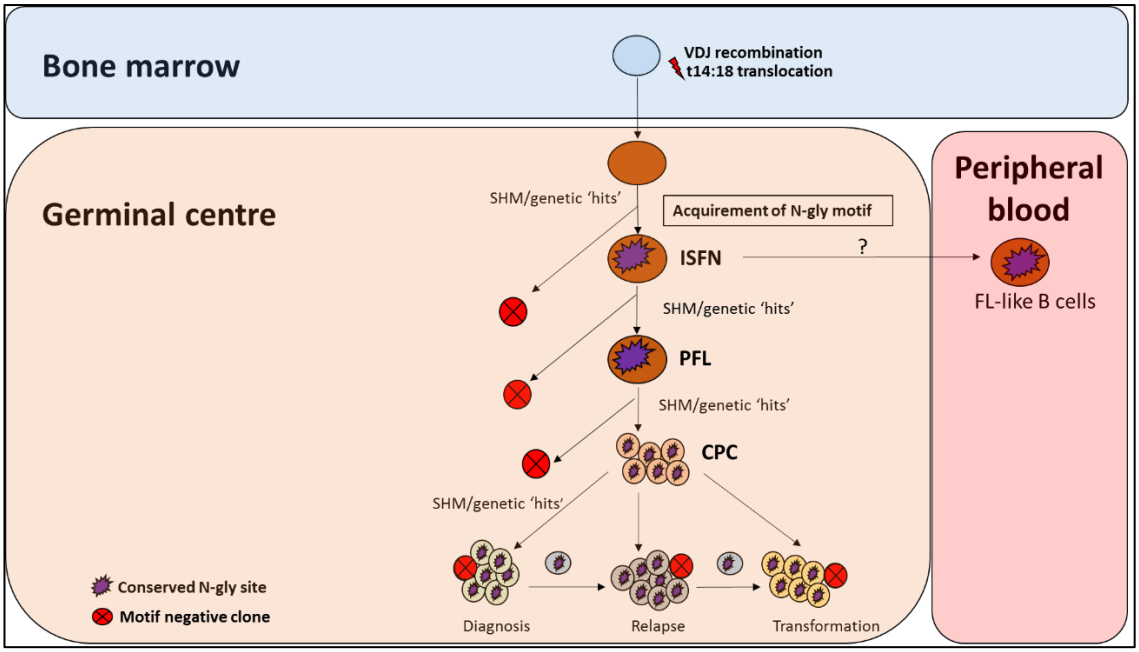


Figure 3.26: *Simplified model of FL evolution and progression based on N-gly motifs.* The model is based on the high throughput sequencing of the IGHV gene. Following the t14;18 translocation during a likely error in VDJ recombination in the bone marrow, the B cell migrates to the germinal centre where it undergoes SHM. N-gly motif sites are acquired early on in the process (purple figure), in the presumed precursor lesion, ISFN. As FL-like B cells are believed to represent the circulating counterparts of ISFN, N-gly motifs may be retained in these cells also. Precursor cells which do not acquire sites through SHM undergo apoptosis, assumed by the low frequency of motif negative subclones in the clonal repertoire. Clones maintain conservation of motifs during tumour evolution through ongoing SHM and gain of additional mutations, leading to an ancestral cell pool population, the CPC. The CPC provides a reservoir from which distinct disease events arise from, all of which retain the motif site. At each stage of evolution, SHM results in the emergence of lost motif clones, due to the largely random nature of the process that does not distinguish between favourable and non-favourable mutations. However, these represent only a minor mass of the heterogeneous tumour population and cannot traffic between events, indicating their insignificance in propagating progression and their likely loss from the clonal repertoire through death pathways. PFL-partial involvement by FL.

Chapter 4: DC-SIGN AND MR EXPRESSION IN FL MICROENVIRONMENT

4.1 Introduction

There is now substantial evidence to suggest that FL cells are able to manipulate and influence their microenvironment to ensure their survival (Ame-Thomas and Tarte, 2014). Compared to healthy follicles, neoplastic GCs have upregulated T cell subsets, such as T_{FH} cells and T_{Regs} cells that were shown to dampen the antitumor response through attenuation of CD8+ cells (Yang et al., 2006a, Yang et al., 2006b). The increase of T_{regs} is likely due to the ability of FL cells to secrete the cytokine CCL22 which is involved in the chemotaxis and migration of T_{Regs} from their interfollicular location into the neoplastic follicle (Yang et al., 2006a).

As previously mentioned, FL cells reside in the GC surrounded by a plethora of non-malignant immune cells, including follicular dendritic cells, and macrophages. A pivotal study by Dave et al., in which the gene expression of non-malignant immune cells was analysed, elucidated how important the immune system, irrespective of tumour genetics, is in determining clinical outcome (Dave et al., 2004). Enrichment for genes expressed in T cells was found to be associated with a favourable prognosis, whereas genes expressed in macrophages and follicular dendritic cells were strongly associated with poor clinical prognosis. Given the highly variable clinical outcome in FL, there is an urgent need to validate these findings to improve patient risk stratification and more importantly, offer the possibility of identifying new targets for future drug design. Using real time PCR and immunohistochemical analysis, Byers *et al* demonstrated that high numbers of CD68+ macrophages and CCR1, the marker of monocyte activation, were associated with shorter survival (Byers et al., 2008). This was supported by Farinha *et al* who also found CD68+ macrophages predicted a worse outcome through immunostaining of FL TMAs (Farinha et al., 2005). The macrophage marker, CD163 that is predominantly found on M2-skewed macrophages (Buechler et al., 2000), has also been associated with poor prognosis and increased angiogenic sprouting in follicular

lymphoma (Clear et al., 2010). While these studies supported those of Dave *et al*, other groups investigating how the number and arrangement of immune cells impacts prognosis has been highly variable, likely due to the lack of conformity regarding treatment and immunohistochemistry scoring techniques (Carreras et al., 2009, de Jong et al., 2009, Lee et al., 2006). Therefore, improved assessment of the microenvironment regarding prognosis is required.

FL cells are believed to interact with their microenvironment through the mannose-lectin interaction, as described in the Introduction. To briefly recap, >90% of FL cases display unusual sites for N-linked glycosylation within their immunoglobulin variable regions that are acceptor sites for high mannose terminating glycans (McCann et al., 2008). Oligomannoses are able to interact with C-type lectins that bind carbohydrate structures in a calcium dependent manner. C-type lectins act as pattern recognition receptors that bind to pathogens and activate immunity (McGreal et al., 2004). Two C-type lectins with specificity for high mannoses are DC-SIGN and MR that are expressed on immune cells including macrophages and dendritic cells (Robinson et al., 2006). These two lectins were shown to bind to oligomannoses expressed on FL surface immunoglobulin and activate BCR signalling, inferring a novel antigen-independent, but cell contact dependent survival pathway for tumour cells.

Interestingly, DC-SIGN and MR have been shown to interact with other cancer-associated glycoproteins, inducing an immune 'dampening' response through upregulation of anti-inflammatory and downregulation of pro-inflammatory cytokines produced by the lectin-expressing cell. Cross-linking of MR on immature dendritic cells with anti-MR antibody was shown to activate DC maturation and induce the secretion of anti-inflammatory cytokines and Th2 attracting chemokines, leading to negative regulation of Th1 polarised responses (Chieppa et al., 2003). Interestingly, similar responses are seen when the receptor is engaged with cancer-derived glycoproteins. For example, ligation of MR with the ovarian cancer derived glycoprotein TAG-72 on TAMs lead to increased IL-10 production and decreased CCL3 (Allavena et al., 2010), the Th1 attracting chemokine. This immune dampening response is also seen when DC-SIGN is engaged with colon cancer derived glycoproteins, leading to an increased IL-10 expression (Nonaka et al., 2008). IL-10 production promotes polarisation of M2

macrophages (Allavena et al., 2010), the phenotype displayed by FL derived TAMs. MR is highly expressed on TAMs and DC-SIGN strongly expressed on M2 macrophages (Amin et al., 2015), suggesting that these lectins play a role in polarising the immune cells to create a pro tumour, immune suppressive microenvironment. Therefore, the mannose-lectin interaction may have a double effect for FL tumourigenesis; through BCR activation and the remodelling of the microenvironment.

DC-SIGN expression has been investigated *in situ* and found to be localised to paracortical lymphatic sinuses and scattered in perifollicular regions (Amin et al., 2015, Linley et al., 2015). Expression was also found on CD68+ mononuclear cells that were shown to be in direct contact with CD20+ FL cells, indicating the feasibility of the mannose-lectin interaction *in vivo*. These studies were performed on fresh frozen tissue sections with either immunofluorescence or immunohistochemistry used to detect antigen expression. Whilst frozen tissue sections provide an excellent source for antigen analysis, information regarding co-expression and spatial relationships between cells is difficult to assess in this type of section, owing to the limitation of multiplex IHC in which a maximum of three antigens can be visualised in a section. A unique technique developed within our department (Ball et al., 2017), has expanded on the multiplex IHC principle by enabling up to seven cellular markers to be explored in a single section, including those which are co-localised through the sequential stripping and re-probing method, described in section 2.27. This offers a more comprehensive and in-depth characterisation of the microenvironment. Stripping and re-probing is performed on formalin fixed paraffin embedded (FFPE) sections. The long-term storage potential of paraffin-embedded tissues provides a wealth of archival samples in which to perform retrospective analysis, often with long-term clinical follow-up available. Furthermore, formalin fixation leads to better preservation of tissue morphology through formation of cross-links between proteins, ensuring observed cellular interactions and tissue location is comparative to that seen *in vivo*.

The quantity of DC-SIGN and MR positive cells in FFPE-derived FL sections has not been investigated, nor their localisation and distribution on different immune cell populations and whether this has any prognostic implications.

4.2 Objectives

We utilised a novel, multiplex, sequential phenotyping strategy to examine the expression and location of DC-SIGN and MR on immune cells of the microenvironment in diagnostic FL biopsies. FL biopsies were taken from two groups at the extremes of overall survival (<5 years and >15 years) to evaluate any prognostic significance. Results were also compared to normal reactive lymph nodes (LNs), in which N-gly motifs are rarely observed in the *IGHV* to determine any correlations between DC-SIGN/MR expression and presence of N-gly motifs.

4.3 Method

4.3.1 Patient samples

Lymph node blocks derived from patients diagnosed with FL at St Bartholomew's Hospital (London, UK) between 1977-2006 were used for this study. 3 µm thick whole sections were cut from paraffin-embedded tissue blocks using a microtome and mounted onto glass slides. Sections were made for nine patients whose survival was >15 years and seven patients whose survival was <5 years. All sections were from diagnostic biopsies. As reactive follicles are considered the healthy counterpart to neoplastic follicles, nine cases of reactive lymph nodes were also sectioned and mounted to be used as a control group. Whole sections were favoured over tissue microarrays (TMAs) to account for the morphologic heterogeneity of FL which can be missed in the 'snapshot' provided by tissue cores consisting of a diameter of 1mm. Clinical data from the extreme survival cases is found in section 4.4.1

4.3.2 DNA extraction

Cores (1mm width by 4mm depth) were taken from follicle rich sites from the extreme survival tissue blocks to obtain tumour cell DNA. DNA was extracted from the paraffin by using the QIAamp DNA FFPE Tissue Kit (Qiagen) and following the protocol. Briefly,

paraffin is first removed from the tissue using xylene followed by removal of the xylene by 100% ethanol which rehydrates the tissue. Proteinase K is a serine protease that was then incubated with the sample to lyse and remove contaminating proteins, including DNAses that degrade DNA and therefore impact integrity. It achieves this by cleaving the peptide bond next to the carboxyl group of hydrophobic amino acid residues. A 90^{o}C incubation reverse formalin crosslinking of nucleic acids before the sample is transferred to the QIAamp MinElute column where DNA binds to the column followed by a series of washes to remove contaminants and is finally eluted. DNA concentration and contaminants were assessed using the Nanodrop Spectrophotometer.

4.3.3 *IGHV* amplification

Eluted genomic DNA then proceeded to *IGHV* amplification as described in Section 3.3.3.

4.3.4 IHC

All slides were first stained for DC-SIGN simultaneously to avoid 'batch effects' with regards to staining intensity. The staining procedure was performed as described in section 2.25. Following coverslipping, slides were incubated overnight in a fume cupboard to ensure complete drying of the DPX mountant. Slides were then scanned using the Pannoramic 250 High Throughput scanner to produce high quality brightfield images. The quality of scanned slides were checked on the Pannoramic Viewer program (Version 1.15.4).

Slides were placed overnight in xylene to remove coverslips before proceeding to stripping and reprobing with the next antibody, as described in section 2.27. Slides were reprobed with antibodies in the following order; CD10, CD163, MR, CD68 and CD21.

4.3.5 Visiopharm

The digital image files were uploaded into the Visiopharm software and the six images for each case were automatically aligned using TissueAlign, followed by manual corrections. The cells of the aligned images underwent analysis for the six markers to

identify cell phenotype. Readouts included csv files documenting the number of single or multi-stained cells with regards to known cell markers CD21 (follicular dendritic cell), CD68 (macrophage), CD163 (M2 macrophage), their location (GC or interfollicular regions) and the mean intensity of DC-SIGN and MR staining in each of these cells. This was performed by the two Visiopharm APPs described in section 2.28.

4.3.6 Statistics

Statistical analyses were performed using GraphPad Prism version 5.4 (GraphPad Software Inc., La Jolla, CA, USA). Statistical analysis was used to compare the different extreme survival groups (<5 year and >15 year) and control group (reactive lymph node). One-way ANOVA without assumption of Gaussian distribution was performed to determine if the differences between the group means were greater than you'd expect to see by chance. Two-way ANOVA was used to determine if the differences between DC-SIGN and MR staining within each group was greater than you'd expect to see by chance. Statistic tests were considered significant at $P \leq 0.05$ (*). A p-value of ≤0.01 was identified by two asterisks (**) and a p-value of ≤0.001 was identified by three asterisks (***).

4.4 Results

4.4.1 N-gly motifs status of different groups

Clinical information regarding our cohort is provided in Table 4.1. The average overall survival of patients within the <5 year survival group was 1.91 years, whereas the overall survival in the >15 year survival group was 22.49 years. The average age of the <5 year survival group was 61 years whereas the average age of the >15 year survival group was 50.5 years. Upon diagnosis, patients underwent either watchful waiting, single chemotherapy or radiotherapy. No clinical information could be gained on the reactive lymph node group except for their absence of a lymphoma diagnosis.

	Patient No.	Sex	Age	FLIPI Score	Initial management	Overall survival (years)
<5 year survival	29	F	54	3-High	Observation	1.38
	30	F	31	2-Intermediate	Observation	2.60
	31	M	72	1-Low	Observation	0.80
	32	F	80	n/a	Observation	3.70
	33	M	73	3-High	n/a	0.43
	34	M	56	n/a	Observation	3.62
	35	F	n/a	3-High	Single chemotherapy	0.88
>15 year survival	36	F	53	1-Low	Radiotherapy	15.27
	37	F	51	1-Low	n/a	25.19
	38	M	37	1-Low	Single chemotherapy	30.54
	39	M	47	1-Low	n/a	25.60
	40	M	62	2-Intermediate	Observation	27.35
	41	M	63	2-Intermediate	n/a	16.25
	42	M	40	3-High	Single chemotherapy	19.01
	43	F	39	1-Low	Radiotherapy	22.55
	44	F	63	3-High	n/a	20.68

Table 4.1: *Clinical information regarding extreme survival patients.*

To determine N-gly motif status within our cohort, *IGHV* amplification and heteroduplex analysis were performed on cores. In 48.5% of cases, I was unable to successfully amplify the *IGHV* gene. 100ng of genomic DNA from these cases was run on a 1.8% agarose gel at 100V to evaluate DNA integrity (Figure 4.1). The lack of distinctive heavy bands and presence of smears was indicative of poor quality, fragmented DNA. To determine whether fragments were similar in size to the *IGHV* product, the housekeeping gene *GAPDH* and *GATA2* were utilised as controls, as they are 200bp and 440bp in length, respectively. Whilst the majority of samples were able to amplify *GADPH*, *GATA2* was not amplified in any cases (Figure 4.1), suggesting the FFPE derived DNA fragments were too short to successfully amplify *IGHV*.

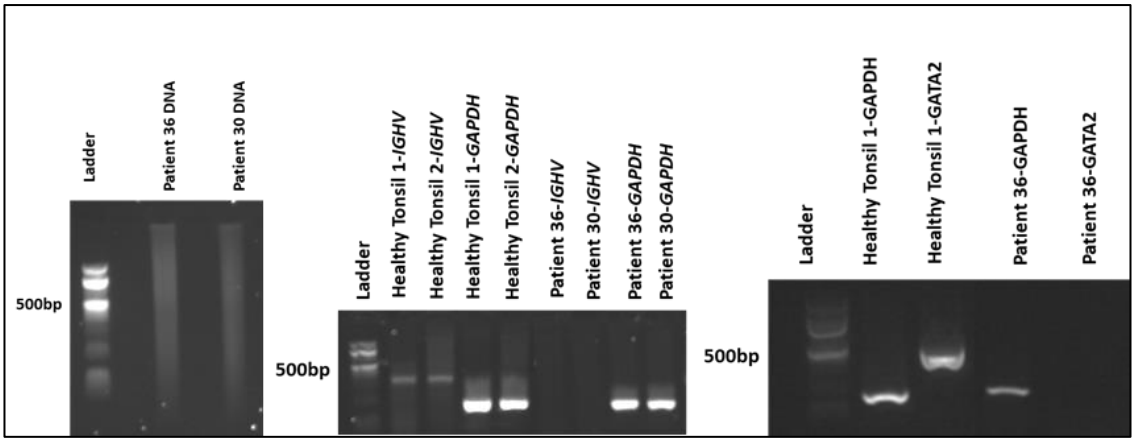


Figure 4.1: *Representative agarose gels highlighting DNA integrity from two extreme FFPE samples*. Left; smear of DNA is indicative of poor quality DNA extracted from two FFPE cases. Middle; DNA extracted from two tonsil biopsies was compared with DNA extracted from FFPE. *IGHV* was successfully amplified in tonsil samples whereas the two FFPE failed amplification. Successful amplification of *GAPDH* (200bp) in both samples is indicative of intact DNA of a certain fragment size. Right; amplification of GATA2, which is a 440bp long product was unsuccessful in patient 36, indicative of DNA fragments being <440bp long. This observation was reproduced in all other samples which failed *IGHV* amplification.

Despite this, the *IGHV* in a number of cases from each group was successfully amplified (Table 4.2). Reactive cases were all N-gly motif negative, keeping in line with the low occurrence seen in normal somatically mutated B cells (~9%) (Alcoceba et al., 2012). The extreme survival groups were all motif positive.

	Patient No.	N-gly motif	Region
Reactive	A	no site	
	B	no site	
	C	no site	
	D	n/a	
	E	no site	
	F	n/a	
	G	no site	
	H	no site	
	I	no site	
<5yr survival group	29	NIT	CDR3
	30	n/a	
	31	NMT	CDR1
	32	n/a	
	33	NFS	CDR2
	34	n/a	
	35	n/a	
>15yr survival group	36	n/a	
	37	NNT	CDR2
	38	n/a	
	39	NMS	CDR2
	40	NFS	CDR2
	41	n/a	
	42	NFS	CDR3
	43	n/a	
	44	n/a	

Table 4.2: *N-gly status in the IGHV of samples undergoing multiplex immunohistochemistry*. n/a indicates cases that failed *IGHV* amplification.

4.4.2 Distribution of immune cells in FL lymph node

Three markers were used to identify three distinct cell types. CD21 is a marker of FDCs, CD68 is a pan macrophage marker and CD163+ is widely seen as a phenotypic marker of M2 macrophages. Usually M2 macrophages are both CD68+ and CD163+, however our analysis revealed that a large population of CD163+ cells were not also positive for CD68. Therefore CD68+CD163+ cells were treated as distinct from CD163+ cells, which were defined as M2 macrophages. Cells which were absent for all three markers but positive for MR or DC-SIGN were labelled 'CD68-CD163CD21-'. The numbers of each cell type were quantified in both the GC and interfollicular regions (termed 'tissue' in the graphs), with actual numbers given in supplementary Table 1.7.

CD21+ cells within the GC were the predominant cell type in the reactive and extreme FL groups. The number of CD21+ cells was significantly lower in the tissue for each group, supporting CD21+ as a robust follicular dendritic cell marker. This was further evidenced by the morphology of CD21+ cells that produced a classical meshwork pattern (Figure 4.2).

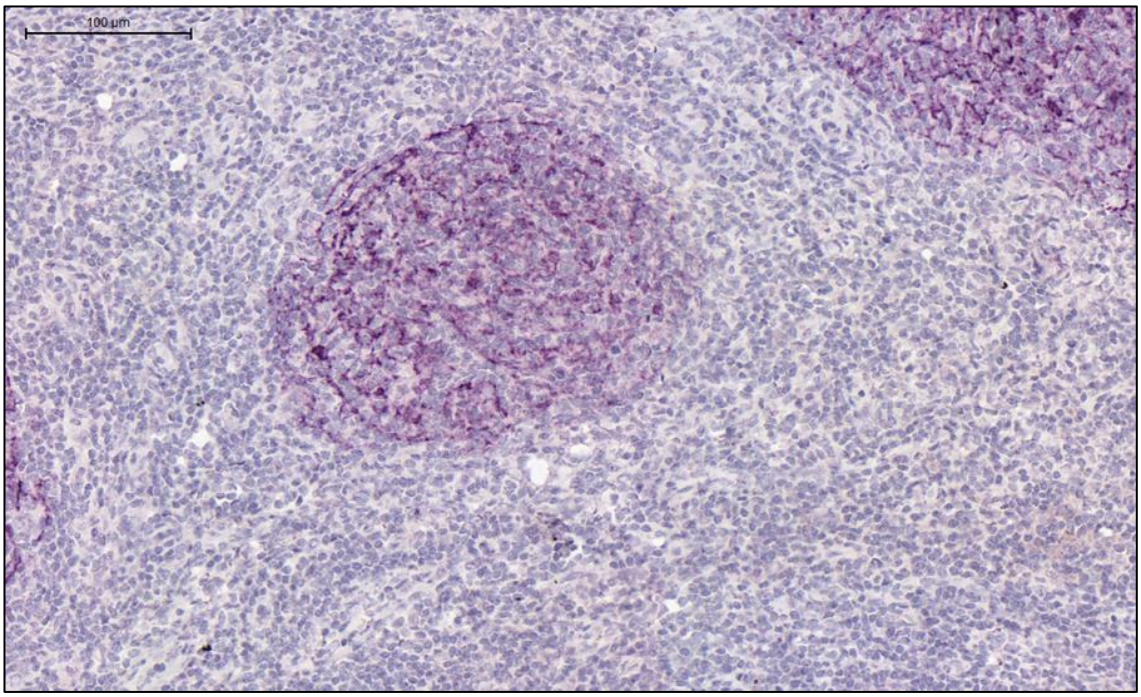


Figure 4.2: *Meshwork pattern of follicular dendritic cells identified by CD21 marker in patient 32.* Scalebar (top-left) represents a distance of 100µm. Magnification: x19.59.

GC derived CD21+ cells made up 32.01% of the overall cell population in the reactive group (Figure 4.3). GC CD21+ cells of the <5 year and >15 year cases were in higher numbers, making up 46.6% and 34.7% of the cell population, respectively. The number of CD68+ in the GC was lowest in the reactive group, making up 0.15% of the overall cell population. <5 year and >15 year groups had comparatively higher CD68+ numbers of 1.4% and 1.77%, respectively. The interfollicular region CD68+ count was highest in the <5 year group. In the GC, the extreme survival groups both had a greater proportion of CD163+ cells compared to the reactive, in contrast to their interfollicular counterparts, in which all three groups had similar cell counts. CD68+CD163+ cell numbers were

slightly raised in the interfollicular regions compared to the GC in all three groups. The reactive group had an increased number of CD68+CD163+ cells in the interfollicular regions compared to the FL groups.

The CD163-CD68-CD21- population within the interfollicular region was the predominant cell type in the reactive and extreme FL groups. This cell group made up 31.9%, 13.4% and 23.5% of the total cell population in the reactive, <5 year and >15 year survival groups, respectively. The reactive LN group had a significantly higher number of these cells compared to the <5 year survival group (*p*-0.05), with a similar trend observed between the reactive and >15 year survival group.

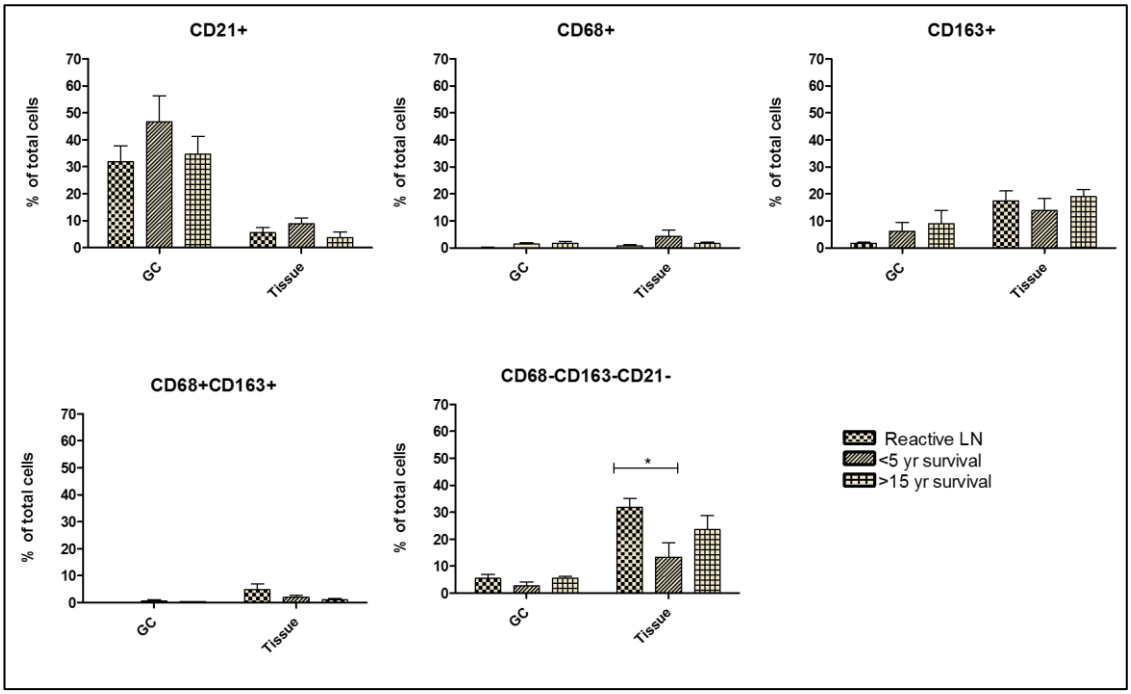


Figure 4.3: *Distribution of cell types in GC and interfollicular regions across all three groups.* The visiopharm software identified positive cells for each marker through staining intensity. The number of positive cells for each marker was divided by the total number of cells identified to give the % values. Total numbers are shown in supplementary Table 1.7. Reactive LN (n=9), < 5 year survival (n=7), > 15 year survival (n=9). * *p*-0.05.

4.4.3 Mean intensity of DC-SIGN and MR

To ensure DC-SIGN and MR staining were specific, I checked by eye the location and distribution of DC-SIGN+ and MR+ cells throughout samples. In FL samples, both stains had a similar distribution pattern within the perifollicular zone and sinusoid, matching what has been reported by Amin *et al* and Linley *et al* in frozen tissue sections (Linley et al., 2015, Amin et al., 2015) (Figure 4.4).

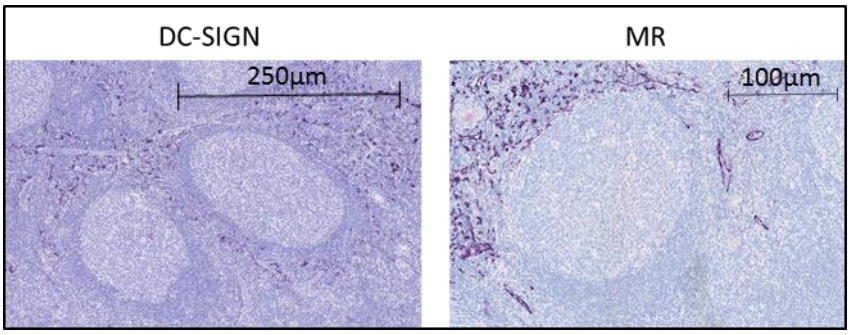


Figure 4.4: *Specific DC-SIGN and MR staining in patient 38.* As reported in literature, positive cells were mostly located to the perifollicular zone. Magnification of DC-SIGN stain: x6.03. Magnification of MR stain: x15.

DC-SIGN and MR staining was assessed for each cell population within each group by calculating mean intensity scores. This was to determine whether DC-SIGN and MR staining differed between specific cell populations (Figure 4.5). DC-SIGN staining was lower than MR staining in both GC and interfollicular regions for all groups. The mean DC-SIGN intensity in GC derived CD21+ cells was significantly higher in the reactive LN group compared to the >15 year survival group (*p*-0.01). <5 year survival also had a lower DC-SIGN intensity compared to the reactive LN group, but without achieving statistical significance. Differences were also seen between the reactive LN group and the FL groups for GC located CD163+ cells and CD68+CD163+ cells, in which the reactive LN group had significantly higher DC-SIGN staining. In contrast, no difference in mean MR intensity was seen in any GC located cell types. Similarly to the GC, DC-SIGN staining of the interfollicular region was greater in the reactive group CD68+ population compared to the >15 year survival group (*p*-0.05), with similar trends seen for the <5 year survival

group. DC-SIGN staining was also greater in the CD163+ population of the reactive group compared to <5 year (*p*-0.01) and >15 year (*p*-0.001) survival groups. MR staining was similar across groups and cell types, excluding CD163+ cells, in which the reactive LN staining was significantly greater than the <5 year FL group (*p*-0.05) and >15 year FL group (*p*-0.001). This is in contrast to DC-SIGN staining in which there are several significant distinctions between the reactive LN and FL groups (described above).

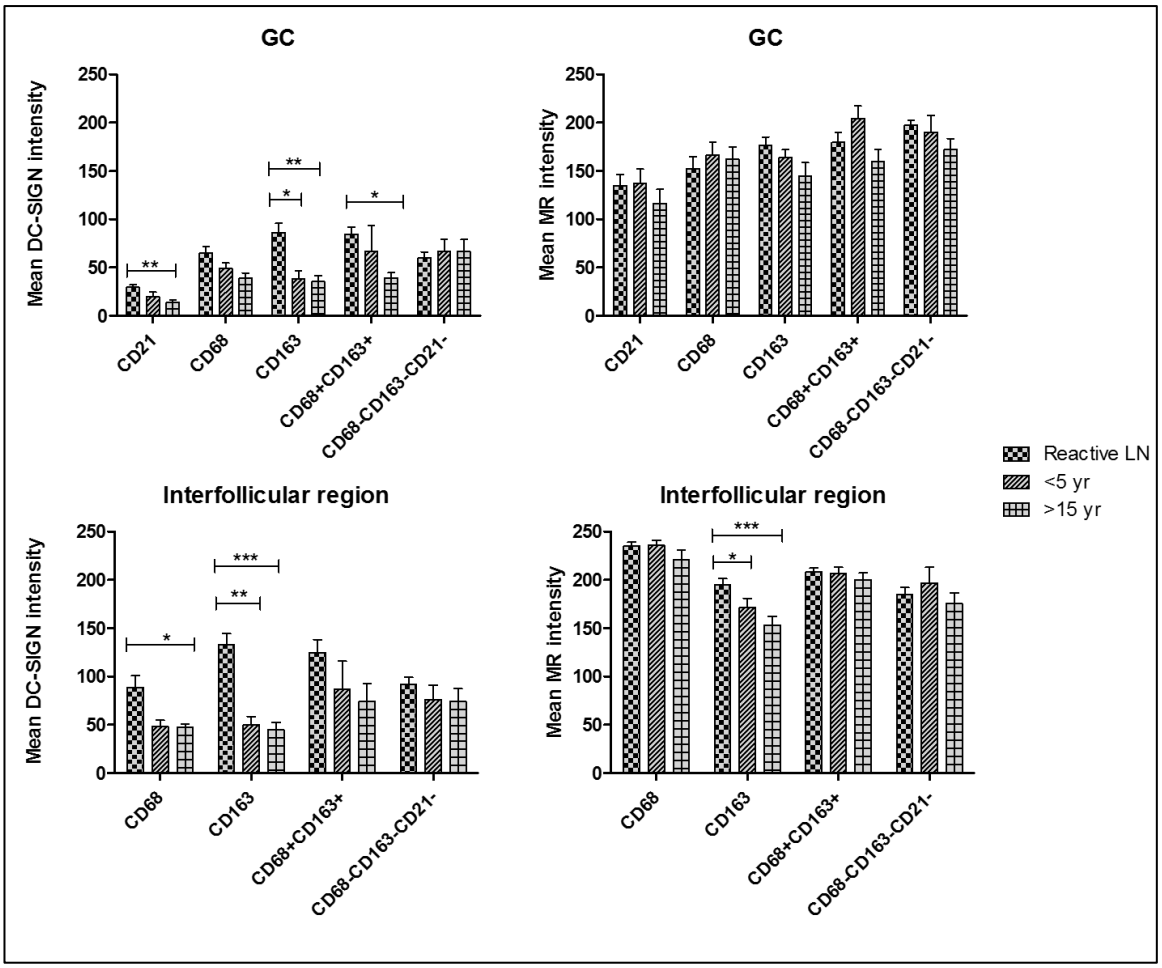


Figure 4.5: *Mean intensity of DC-SIGN and MR across cell types and disease groups.* For each cell within a cell type group, a DC-SIGN and MR intensity value was given. The mean intensity value for all cells within a cell type group was calculated and plotted.

4.4.4 DC-SIGN and MR markers in GC-located cell populations

Individual cells within the GC were assessed for DC-SIGN and MR positivity by comparing their individual intensity values to the median DC-SIGN and MR intensities calculated across all cell types. This was done on an individual case basis to determine inter-patient heterogeneity. Cells with intensity greater than the median value were termed positive whereas cells with intensity lower than the median value were termed negative. This was performed in Excel (Microsoft) using the COUNTIF function. The results for CD21, CD68 and CD163 positive cells in the individual cases can be seen in Figure 4.6. CD68+CD163+ and CD68-CD163-CD21- were omitted from analysis due to these cell types not being identified in some patients. From Figure 4.6, we can observe the heterogeneity of DC-SIGN and MR expression across different cell types, along with intra-group heterogeneity of expression.

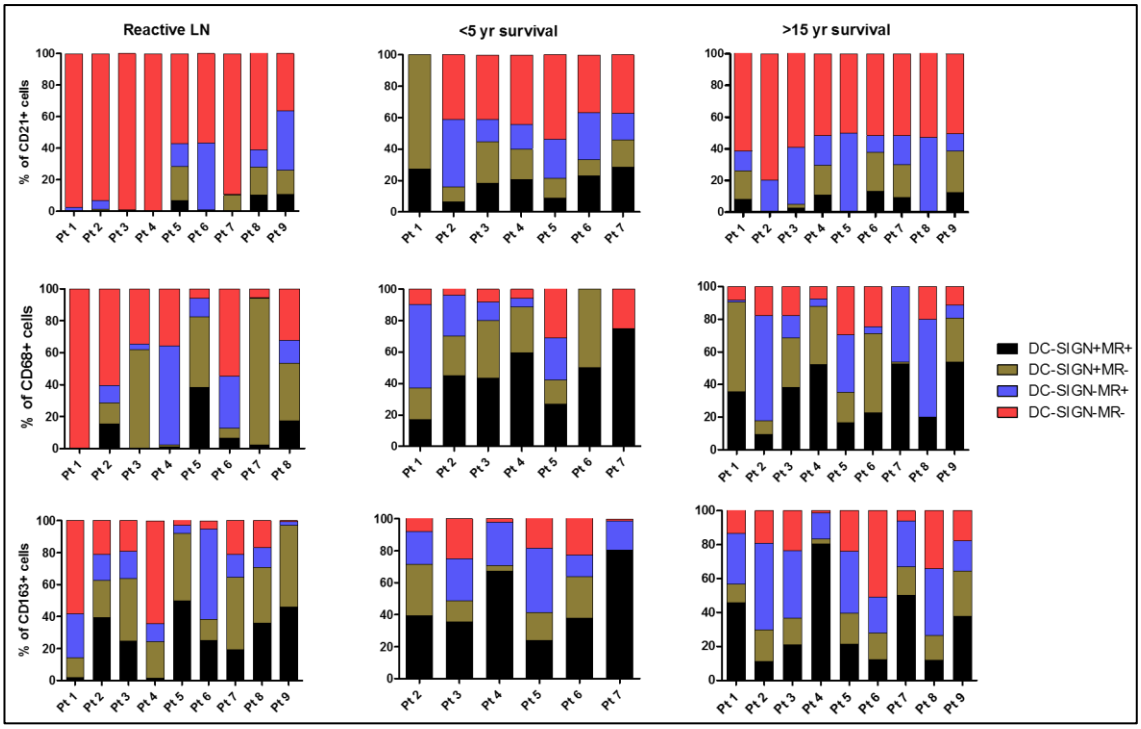


Figure 4.6: *Distribution of immune cells with DC-SIGN and MR positive and negative phenotypes within individual cases.* Top panel: % of CD21+ cells across groups. Middle panel: % of CD68+ cells across groups. Bottom panel: % of CD163+ cells across groups. Proportion of double positive (DC-SIGN+MR+) cells are represented in black, proportion of single DC-SIGN positive cells are represented in brown (DC-SIGN+MR-), proportion of single MR positive cell are represented in blue (DC-SIGN-MR+) and the proportion of double negative (DC-SIGN-MR-) are represented in red. These bar charts highlight inter and intra-group patient heterogeneity.

I then compared whether there was any distinction between DC-SIGN and MR expression within cell populations of different groups (Figure 4.7). A larger proportion of CD21+ cells in the <5 year survival were double positive compared to the reactive group (*p*-0.01). This was also the case for DC-SIGN expressing CD21+ cells. The double negative population was significantly lower in the <5 year group compared to the reactive group. The phenotype of CD21+ cells in the >15 year also showed a similar trend to that of the <5 year group but to a lesser degree. For MR positive cells, the FL groups all had greater proportions compared to the reactive group.

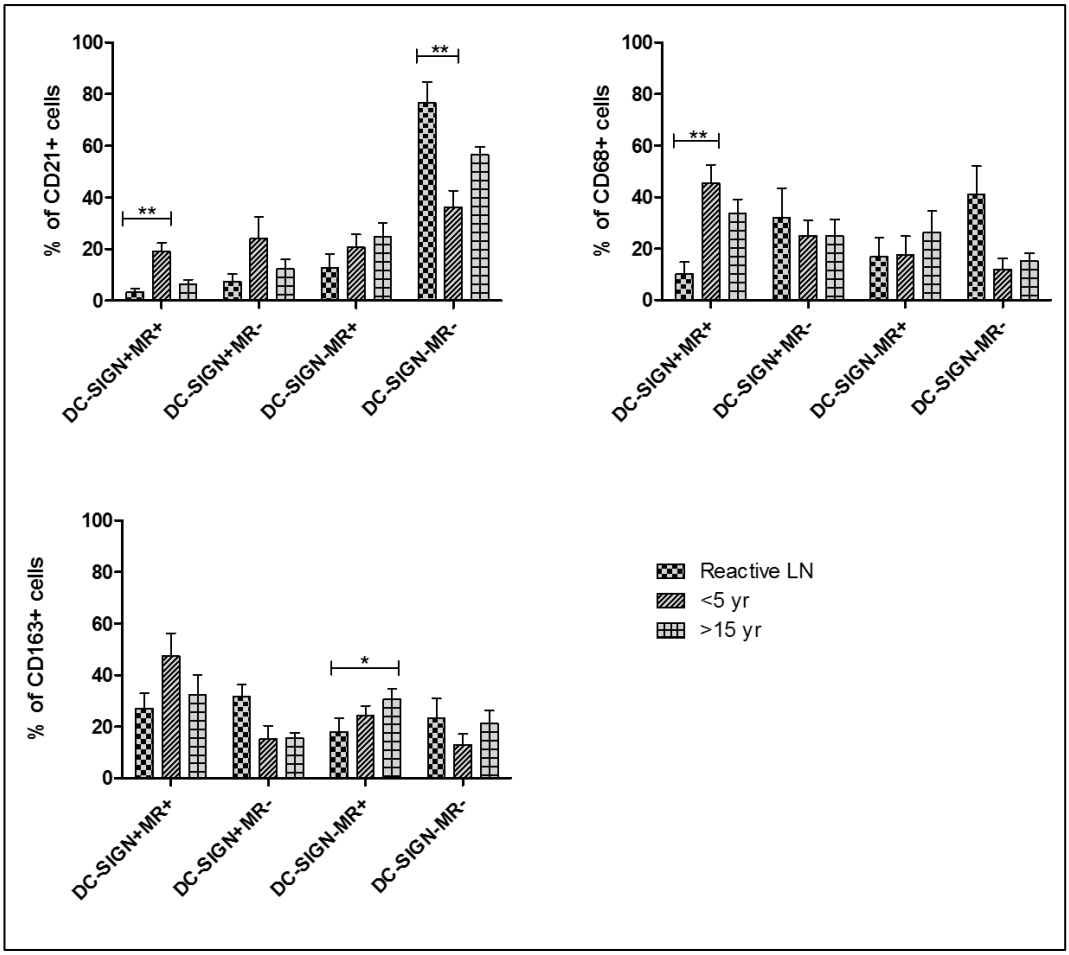


Figure 4.7: *% of GC located cell types expressing DC-SIGN and MR positive and negative phenotypes across disease groups*. This figure is a re-representation of the data in Figure 4.6, with the mean values for each group plotted. * p-0.05, **p-0.01.

For the FL groups, the double positive phenotype made up the largest proportion of CD68+ cells whereas for the reactive group, this phenotype was the least represented phenotype, in contrast to its tissue counterpart. Indeed there was a significant difference between the reactive LN and <5 year survival group (*p*-0.01) with similar trends observed for the other FL groups. Similarly, to CD21+ cells, the double negative population was lower in the FL groups compared to the reactive LN group in which this cell population represented the predominant phenotype of CD68+ cells.

For the CD163+ population, the FL groups had a greater proportion of double positive cells compared to the reactive, with the biggest proportions seen in the <5 year group. The DC-SIGN population was greatest in the reactive group compared to the FL groups.

For the double negative group, the reactive and >15 year groups had similar proportions whereas the <5 year group was lower. However the FL groups having a greater proportion of MR+ cells, with the difference most obvious between the reactive and >15 year survival groups (*p*-0.05).

4.4.5 DC-SIGN and MR markers in interfollicular-located cell populations

DC-SIGN and MR positivity was then assessed in cell populations within the interfollicular region. Figure 4.8 shows the % of different cell types in the tissue which were either double positive for DC-SIGN and MR, single positive for either DC-SIGN or MR, or negative for both markers. Similarly to the GC, the predominant phenotype of CD21+ cells across groups was double negative. For CD21+ cells, a greater proportion of cells were double negative in the reactive LN group compared to the three FL groups, in contrast to the % of MR+ cells which was lower in the reactive LN group compared to the FL groups (Figure 4.8). % of double positive cells was higher in the FL groups, with similar trends seen for DC-SIGN positive cells, although findings were not significant. The double negative was the dominant phenotype of CD21+ cells across groups in the tissue. This is in contrast to the CD68+ population, in which the dominant phenotype was double positive followed by MR+ cells. DC-SIGN+ and double negative phenotypes made up the lowest % of CD68+ cells. Compared to CD68+ cells in the GC, there was a lower proportion of DC-SIGN positive and double negative GC located CD68+ cells across all groups.

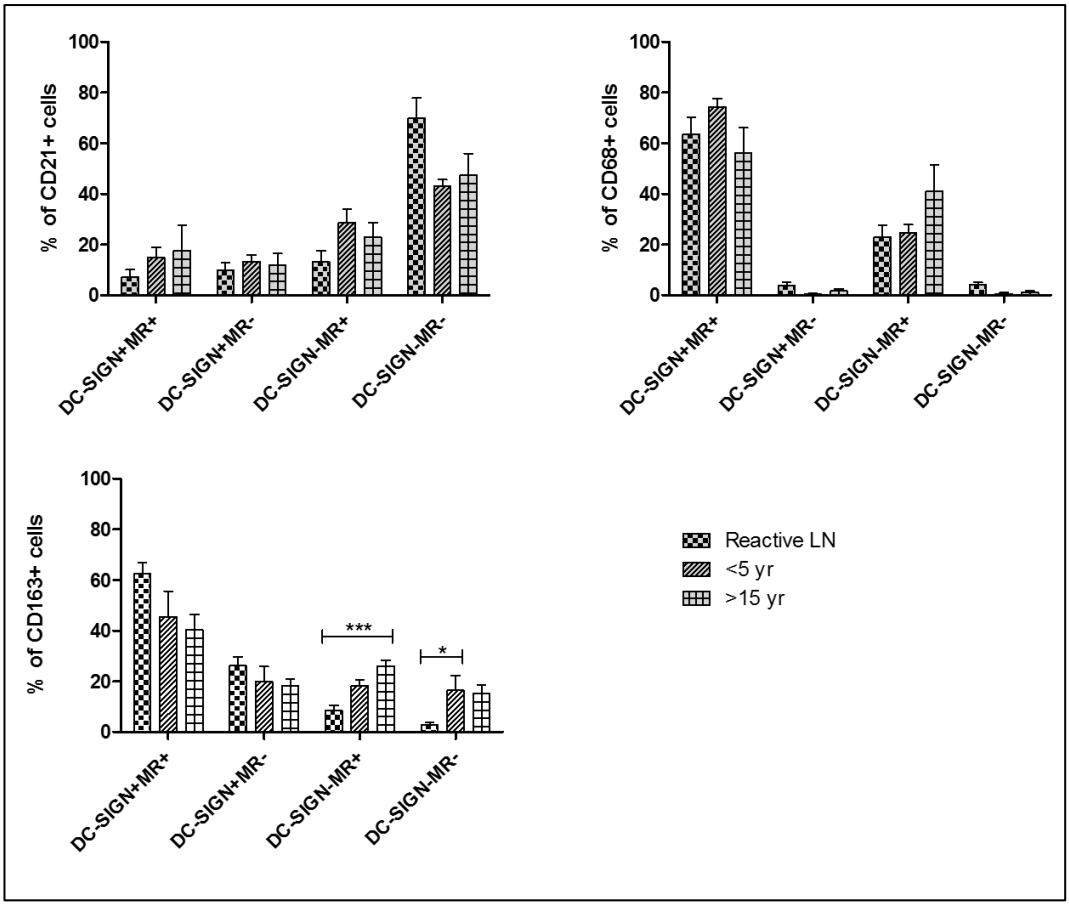


Figure 4.8: *% of interfollicular located cell types expressing DC-SIGN and MR positive and negative phenotypes across disease groups.* The mean values for each group is plotted. * *p*-0.05, *** *p*-0.001.

For CD163+ cells, the double positive phenotype was dominant across groups, with reactive LN having relatively high cell counts compared to <5 year and >15 year survival groups. Reactive LN group had a significantly lower expression of double negative CD163+ cells compared to the two extreme survival groups (*p*-0.05). Similar to the trend seen in the GC counterpart, the DC-SIGN population was greatest in the reactive group compared to the FL groups

The proportion of MR+ cells was significantly higher in the >15 year group (*p*-0.001) with the 5 year survival also showing an upward trend (Figure 4.8).

4.5 Discussion

To our knowledge, this is the first study that has utilised a unique and in-depth sequential multiplex procedure to comprehensively analyse the FL microenvironment of extreme survival cases to determine prognostic impact. So far, traditional multiplex IHC studies regarding the assessment of the tumour landscape have been limited by the restricted number of antigens assessed within a section due to antibody cross reactivity (detailed in section 2.26) along with the inability to identify antigen co-localisation. Through a stripping and reprobing method on single FFPE whole sections, we have been able to; 1) assess the number and localisation of several immune cells within the FL LN, 2) determine and quantify immune cells expressing DC-SIGN and MR, and 3) assess microenvironment distinctions between healthy and disease states as well as the differences within extreme FL survival groups.

Reactive LNs were used as a control group as they represent the healthy counterparts of neoplastic follicles. Furthermore, as normal somatically mutated B cells rarely express N-gly motifs in their *IGHV*, they offered the benefit of determining whether any differences observed regarding DC-SIGN and MR expression between healthy and disease tissue could be correlated to N-gly status. To determine N-gly status, we extracted genomic DNA derived from the FFPE samples and amplified the *IGHV* gene. In 44% of cases, we were unable to successfully amplify the gene. This was likely due to the quality of the DNA as indicated in Figure 4.1. DNA extraction from FFPE is associated with a number of challenges, mainly due to formalin fixation that results in cross-links between proteins and DNA. This causes the DNA to be significantly degraded and fragmented. The degree of fragmentation is also influenced by the age of the sample and as the majority of our extreme survival group biopsies were taken during the 1970s and 1980s, this may be a valid reason for the PCR failure (Figure 4.1). Despite this, the *IGHV* was successfully amplified in a number of cases across the groups. The reactive LN cases were motif negative whilst our extreme survival groups were positive, supporting previous findings (McCann et al., 2008, Alcoceba et al., 2012).

Following multiple IHC and identification of different cell types, I assessed the distribution of cell types across the GC and interfollicular regions. Usually M2

macrophages are both CD68+ and CD163+, however our analysis revealed that a large population of CD163+ cells were not positive for CD68. This could be due to a weaker staining by CD68, which has been noted in other studies (Kridel et al., 2015b, Tan et al., 2012, Klein et al., 2014). Therefore CD68+CD163+ cells were treated as distinct from CD163+ cells, which were considered M2 macrophages.

FDCs (CD21+) made up the largest proportion of cells within the GC across groups, with the <5 year group having the highest number, followed by the >15 year group. FDC associated genes were shown to correlate with a poor outcome by Dave *et al*, which supports our findings (Dave et al., 2004). In the interfollicular region, M2 macrophages made up the largest proportion of cells for the FL groups, with similar counts seen for the reactive LN. However, when we compared M2 macrophages in the GC, the FL groups had a higher number compared to the reactive group, indicative of tumour-associated macrophages (TAMs). The prognostic implications of the CD163 marker has been explored by Kridel *et al.* Analysis of two studies revealed that CD163 was predictive of outcome in FL but was dependent on the type of treatment patients had received (Kridel et al., 2015b). PRIMA trial patients treated with rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone, and randomized to rituximab maintenance or observation, had favourable outcome with high CD163+ counts. In contrast, high CD163+ counts in patients treated with first-line systemic treatment including rituximab, cyclophosphamide, vincristine, and prednisone were predicted to have an adverse outcome. No studies have so far looked at CD163 count in diagnostic samples before treatment and therefore our data gives a unique insight into the 'natural' microenvironmental landscape and how this translates to OS. The >15 year survival group had a higher CD163+ count in the GC compared to the <5 year, suggesting a favourable outcome of higher CD163+ counts in disease, however the difference in count between the two was not significant.

The mean intensity of DC-SIGN and MR was assessed between groups. This was first assessed across all cell types. In all groups, MR intensity was significantly higher than DC-SIGN intensity. This could be indicative of a greater expression of MR compared to DC-SIGN within the lymph node or a comparatively weaker signal of DC-SIGN expression due to the antibody. MR intensities were similar between groups except for the

interfollicular located M2 population, in which the mean intensity for the reactive group was significantly higher than the >15 year group. For DC-SIGN, there was intergroup variability in GC located follicular dendritic cells, M2 macrophages and CD68+CD163+ cell populations in which the reactive LN group had a significantly higher mean intensity. This was also true for DC-SIGN staining of the macrophage and M2 populations in the tissue, suggestive of greater DC-SIGN expression in the reactive group compared to the disease groups. However, the extreme survival tissue blocks were considerably older than the reactive LN blocks, which could have influenced staining intensity. Indeed Grillo *et al* found that membrane antigens presented reduced staining intensity in older blocks (Grillo et al., 2017). Therefore, it was important to determine on a case-specific basis, the degree of staining of DC-SIGN and MR to determine the number of cells that were either DC-SIGN or MR positive, double positive or double negative. This was achieved by calculating the median DC-SIGN and MR intensity values and grouping individual cells into either positive or negative categories based on whether they exceeded or fell below the median values.

DC-SIGN and MR were expressed on a proportion of macrophages, M2 macrophages and FDCs across groups, which is an expected finding given their normal expression on these cells types. They have a variety of roles as pattern recognition receptors, including phagocytosis/antigen presentation of pathogens and mediation of endogenous cell-cell interaction during the innate immune response (Dambuza and Brown, 2015, Geijtenbeek and Gringhuis, 2009). DC-SIGN+MR+ was the predominant phenotype of both macrophages and M2 macrophages of the interfollicular region in all groups.

In the GC, the phenotype of macrophages and M2 macrophages was more evenly distributed. However, the <5 year survival group had a significantly greater proportion of double positive macrophages compared to the reactive LN group (p-0.01). The >15 year survival group also had a higher proportion of double positive macrophages compared to the reactive group, but did not reach significance. A similar trend for both FL groups was seen in the other cell types, regarding higher numbers of double positive cells compared to the reactive group. In the FDC compartment, <5 year survival group had a significantly higher number of double positive cells (p-0.01). These differences between disease and reactive groups were not seen in the interfollicular region cell

counterparts, indicative of skewing of GC-specific macrophages, M2 macrophages and FDCs to display both DC-SIGN and MR. These higher numbers inversely correlated to fewer double negative cells compared to the reactive LN group, with a significant difference seen with the <5 year survival group (*p*-0.01), suggesting a skewing of immune cells to the other three phenotypes in disease.

M2 macrophages derived from the >15 year survival group had significantly higher numbers of the MR positive phenotype compared to the reactive group in both the GC (*p*-0.05) and interfollicular region (*p*-0.001), with the <5 year survival group showing a similar trend without significance. The higher count of MR positive M2 cells in the interfollicular region may indicate that these calls can exert pathological effects without being in direct contact with tumour cells.

Compared to the reactive group, the extreme survival groups in general shared similar trends concerning the DC-SIGN and MR phenotype of immune cells. Between the two groups, no significant difference was reached. However, this was a small cohort and therefore analysis of more patients is needed to validate DC-SIGN or MR expression as predictive markers of outcome. Studies by Dave *et al* and others have revealed the impact of immune cells in patient outcome, with CD68+ macrophages predictive of a worse outcome (Dave et al., 2004, Farinha et al., 2005, Byers et al., 2008). CD163+ M2 macrophages have also been associated with a poorer clinical outcome, however Kridel *et al* demonstrated that this prediction is heavily impacted by treatment (Kridel et al., 2015b). Therefore going forward, we would ideally perform analysis on extreme survival cases that have all undergone the same therapy, to ensure intrinsic biological heterogeneity of the disease is the cause of distinctive clinical outcomes and is not down to treatment heterogeneity. It would also prove interesting to explore lectin expression in histologically and clinically distinct FL subtypes, such as PTFL. These patients usually presenting with localised stage one lymphadenopathy which has a good response rate to local excision or minimal chemotherapy (Araf and Fitzgibbon, 2016). Compared to adult FL, it is genetically 'light', with fewer recurrent mutations and lack of epigenetic mutations, indicative of a distinct pathogenesis pathway (Louissaint et al., 2016). Interestingly a group recently observed that a further distinction between adult FL and PTFL was a lack of N-gly motifs in the *IGHV* gene (unpublished data), suggesting that

PTFL does not utilise the mannose-lectin interaction. As we have access to PTFL sections from our collaborators in Germany, Tubingen, this will form part of our future investigation.

The distinctions between reactive and FL LNs suggests that presence of N-gly motifs in the *IGHV* are associated with a skewing of immune cells to express lectins that are implicated to be involved in the mannose-lectin interaction. Indeed, the majority of increased DC-SIGN and MR cells are seen in the GC compartment, where FL cells reside. Therefore, we can infer that through this interaction, lymphoma cells can remodel their microenvironment by influencing the lectin-expressing cells. DC-SIGN and MR interaction with other cancer derived glycoproteins has been shown to contribute to a pro-tumour microenvironment (as discussed at the beginning of this chapter), indicating the feasibility of this model. A greater proportion of lectin-expressing cells may ensure sustained BCR activation and survival of the tumour cell within the GC.

This study has shown an in-depth characterisation of the FL microenvironment and has revealed several significant differences between the expression of DC-SIGN and MR on immune cells within disease and healthy reactive LNs. These preliminary findings support further in-depth analysis using greater cohort numbers. Whilst we have demonstrated the promise of sequential multiplex IHC in analysing the microenvironment at a single cell level, mass cytometry imaging is an exciting new method of contextualising the landscape. Up to 37 antigens can be visualised at once, giving unprecedented insight into the microenvironment. This will be advantageous in exploring the precise spatial arrangement of lectin-expressing cells in the interfollicular and GC regions and their interaction with FL cells and immune cells, including T cells to provide a comprehensive view of the mannose-lectin interaction *in situ.*

Chapter 5: DISCUSSION

FL is the most prevalent indolent NHL in western countries. The introduction of the anti-CD20 monoclonal antibody rituximab revolutionised FL treatment by increasing response rates and significantly improving PFS and OS (Hiddemann et al., 2005, Marcus et al., 2005). Patients now have a median survival of 15 years (Tan et al., 2013, Czuczman et al., 2004). Despite this significant breakthrough, and patients responding well to initial treatment, the disease remains incurable with patients undergoing frequent relapses with each remission free period becoming progressively shorter before patients become refractory to treatment or transform to an aggressive lymphoma. The clinical burden due to the protracted relapse-remitting course is significant and therefore better understanding of key pathogenic pathways is required to develop targeted therapies.

A significant challenge to deciphering this is the heterogeneity of FL disease, both clinically and biologically. The majority of patients achieve a high response rate to first-line therapy and usually follow an indolent disease course. However, in up to 25% of patients, progression occurs within 2 years of treatment or the disease transforms to a high-grade lymphoma (Casulo et al., 2015, Casulo, 2016). For these patients, prognostic outlook is relatively very poor and predicting which patients would fall into this high-risk group remains an ongoing area of research (Huet et al., 2018b) as it would identify patients most likely to benefit from clinical trials. The biological heterogeneity of FL has been exemplified by NGS that has revealed a complex and variable genetic landscape. Recurrent mutations in specific genes have been identified, indicating biological pathways perturbed in disease, including epigenetic regulation (Okosun et al., 2014, Araf et al., 2016). However, these mutations are not found within all patients and vary in clonal dominance, not just between patients but also in spatially separated biopsies from individual cases (Araf et al., 2018). Targeting a mutation in a clonal population may, therefore, enable a minor subclone to gain dominance, leading to relapse. This heterogeneity of subclones and their genetics compromises the approach of precision medicine. Targeting early events in pathogenesis which are conserved clonally would likely have the most beneficial impact on sustained remission, which is why

characterising events of the putative CPC compartment is an important new area of research in FL. Genetic changes in the CPC are being investigated and characterised (Okosun et al., 2014), yet CPC-microenvironment interactions remain unexplored despite our extensive knowledge on the importance of microenvironment-tumour cell interactions in pathogenesis.

We implemented a NGS approach to identify and catalogue N-gly motifs in the *IGHV* of unique subclones that contribute to the heterogeneous clonal repertoire of disease. A longitudinal profiling strategy using paired FL and tFL samples enabled the delineation of the clonal *IGHV* repertoire throughout a patient's disease course, giving insight into the importance of motifs during disease evolution, expanding our views on the ontogeny and pathogenesis of FL. The inclusion of a clinically and biologically variable patient cohort ensured findings were representative of the heterogeneous disease. A striking finding was that >97% of subclones within a disease event shared the same motif site as that identified in the major clone, indicating N-gly motifs as a feature of the clonal repertoire. These motif sites were conserved in the clonal population of successive disease events, including relapse and transformation, indicating that motifs were acquired at an early time point of disease evolution in an ancestral cell population from which disease events arise from, the CPC. This is an important finding, as the CPC is believed to be currently resistant to therapy as highlighted by the high relapse rates seen for patients, indicating the therapeutic promise of targeting specific characteristics within this compartment. We've shown motifs to be important in both 'rich' and 'sparse' evolutionary models, suggesting that targeting of these motifs may have universal benefit to all FL patients, regardless of genetic heterogeneity amongst tumour clones. These findings were not restricted to our own experimental data, but also *IGHV* sequencing data generated from different platforms and research institutes, indicating the validity and reproducibility of our data.

The early acquirement of N-gly motifs has previously been demonstrated in study of ISFN cells, believed to represent malignant precursor cells (Kosmidis et al., 2017, Mamessier et al., 2015). Having a functional mannose-lectin interaction at this early stage in disease development may explain how precursor cells are able to survive in the hostile GC without the need for high-affinity BCRs, enabling accumulation of genetic hits

required for malignant transformation. Whether or not motifs are functional within tFL disease is questionable. Following transformation, the disease changes dramatically and becomes aggressive with acquirement of new mutations such as *MYC*. Therefore microenvironmental interactions before transformation may become obsolete and therefore the presence of motifs within subclones specific to this disease event may, in fact, be an 'imprint' of the ancestral cell from which they're derived. However, the striking finding regarding the lack of motif negative subclones passed onto successive disease events, including transformed events, is a strong argument that subclones have an ongoing reliance on motif presentation as the disease progresses. The fate of subclones is determined by N-gly motif status, regardless of the genetic profile of the clone, highlighting the importance of microenvironmental interactions in disease progression. Targeting motifs or the mannose-lectin interaction may lead to significant disruption of the microenvironment and tumour cell death. As a feature of the putative CPC, therapeutic targeting may also lead to longer remission periods and even cure by disrupting the reservoir pool that propagates the disease. Determining the genetic profile of these motif negative clones may provide us with insight into the relationship between genetics and the microenvironment that determines clone fate during disease progression and thereby piece together the profile of progression-associated subclones. It could also give an indication into true FL 'driver' mutations. 'Driver' mutations in this context represent critical mutations that are critical for malignant transformation and disease progression, in contrast to so-called 'passenger' mutations that are assumed to be biologically neutral and simply carried along with the drivers (Stratton et al., 2009). Examples of driver mutations in other haematological diseases include *MYC* in Burkitt's lymphoma and *BCL-ABL1* in chronic myeloid leukaemia (Koschmieder et al., 2005, Schmitz et al., 2014). The t14;18 translocation is considered a 'driver' mutation in FL with the bcl2 mouse models developing follicular hyperplasia (McDonnell and Korsmeyer, 1991, Egle et al., 2004). However, the inability of these models to faithfully emulate the exact indolent nature of the disease and the presence of t14;18 positive cells in healthy people (Limpens et al., 1995, Dolken et al., 1996, Schuler et al., 2009), suggest additional 'driver' mutations are required for malignant transformation. Okosun *et al* revealed through WGS of paired FL-tFL samples, that *CREBBP* and *KMT2D* mutations are likely to be early 'driver' events due to their clonal dominance across temporal samples and

recurrence across patients (Okosun et al., 2014). The 'driver' potential of these mutations is currently undergoing investigation in transgenic mouse models to determine their role in disease initiation. Patients 23-25 were included in this study, with both epigenetic genes considered early 'driver' events. However, the role of these mutations in 'driving' disease progression has been challenged by our study. If these mutations were considered truly clonal, we would assume each subclone to contain the aberrations. The disappearance of motif negative subclones in subsequent samples, with new motif negative subclones maintaining an extremely low count in progression events, implies that subclone survival is more weighted by microenvironmental interactions rather than these 'driver' mutations. Alternatively, it could be the result of motif negative subclones being outcompeted by motif positive subclones due to the advantages of additional microenvironmental support, mediated by the mannose-lectin interaction. However, we also cannot rule out that these motif negative clones have a different genetic profile to motif positive subclones in which driver mutations are absent. As 10% of FL cases are negative for N-gly motifs, it would be highly valuable to determine and compare their genetic profiles with motif positive cases to identify true 'driver' mutations that seemingly make the putative mannose-lectin interaction redundant for disease initiation and progression. The collaboration between two distinct yet interconnected disciplines of FL research could lead to a comprehensive and well-rounded understanding of the intricate pathobiology of FL.

We next utilised a novel multiplex, sequential phenotyping of FFPE sections derived from extreme survival cases of FL (Ball et al., 2017). We were able to look at six different immune cell markers to determine the expression of DC-SIGN and MR on immune cells of the tumour microenvironment and whether this correlated to prognostic outcome. To our knowledge, this is the first study of its kind giving novel insight and characterisation of the microenvironment at unrivalled depth due to a unique stripping and re-probing protocol. We observed a significant deviation in the tumour microenvironment, characterised by a higher proportion of follicular dendritic cells, macrophages and M2 macrophages displaying a MR+ and DC-SIGN+MR+ phenotype compared to the immune cells of reactive LNs, suggesting a pathogenic-derived cause to increased lectin expression within these cells. As reactive LNs were motif negative,

this increase could be attributed to the mannose-lectin interaction that may result in polarisation of interacting immune cells to express these phenotypes. Although there were no significant differences between the extreme survival groups, this data provides the rationale for conducting analysis in larger cohorts to validate the prognostic implications of DC-SIGN and MR expression and provide insight into whether targeting such immune changes can divert the FL microenvironment from one that is tumour supportive to one that is tumour suppressive. It also warrants further investigation into the mannose-lectin interaction *in situ*, which we will explore through high-throughput multiplexing platforms, including mass cytometry imaging.

The knowledge that FL is a significantly heterogeneous disease warrants research into finding key pathogenic events that are universally perturbed in all patients to ensure maximal and widespread benefit. This project has contributed to this aim by determining N-gly motifs in the *IGHV* are an early event in disease evolution of the majority of FL patients irrespective of the genetic and biological heterogeneity of disease. This finding supports the hypothesis of the mannose-lectin interaction operating a critical survival-signalling pathway in FL cells and provides strong rationale for future investigations into translating this finding into improving the outcome for FL patients.

Supplementary Tables

Sample ID	*V* gene	Homology (%)	No. of sites	Region	aa Motif
A	V3-11	89.8	1	CDR2	NIS
B	V3-48	87.3	1	CDR2	NIS
C	V3-23	90.1	1	CDR1	NIT
D	V3-15	87	2	CDR1, CDR3	NFS, NIT
E	V3-15	86.7	1	CDR2	NIS
F	V3-11	82.11	1	CDR2	NIS
G	V3-11	89.5	1	CDR3	NFS
H	V3-15	92.5	1	CDR2	NKS
I	V3-48	87	1	CDR2	NMS
J	V4-59	85.5	1	CDR2	NIS
K	V3-48	89	1	CDR1	NIS
L	V3-30	89.3	1	CDR2	NVS
M	V3-23	88.2	1	CDR3	NIS
N	V3-30	89.7	1	CDR2	NKS
O	V3-11	90	1	CDR2	NIT
P	V3-23	90.2	2	CDR1, CDR2	NFS, NVS
Q	V3-11	90.7	1	CDR2	NIS
R	V3-23	84.8	2	CDR2, FR3	NST, NIS
S	V3-23	87.6	2	CDR1, FR3	NFS, NIS
T	V3-11	88.9	1	CDR3	NIS

Table 1.1: *Incidence of novel N-glycosylation sites in extension cohort of diagnostic FL samples.*

	Barcode sequence	PF Clusters	Yield (Mbases)	% >= Q30 Bases	Mean Quality Score
Patient 23-Diagnosis	CGGCATTA	1,320,495	663	88.79	35.22
Patient 23-Transformation	CACGCAAT	1,319,287	662	90.75	35.65
Patient 24-1st relapse	GGAATGTC	1,188,214	596	90.06	35.49
Patient 24-3rd relapse	TGGTGAAG	944,471	474	90.12	35.49
Patient 24-Transformation	GGACATCA	1,546,380	776	89.37	35.32
Patient 25-2nd relapse	GGTGTACA	929,244	466	87.89	35.02
Patient 25-3rd relapse	GATAGCCA	1,224,512	615	89.03	35.27
Patient 25-Transformation	CCACAACA	1,365,431	685	89.14	35.26

Table 1.2: *Summary of NGS data generated for each patient sample.*

	Merged Reads	Not Merged Reads	Total Reads	Merged Over Total
Patient 24-1st relapse	1,169,453	73,136	1,242,589	94.11%
Patient 23-Diagnosis	1,192,523	52,466	1,244,989	95.79%
Patient 25-2nd relapse	1,046,604	43,035	1,089,639	96.05%
Patient 25-3rd relapse	829,744	38,671	868,415	95.55%
Patient 25-Transformation	1,375,958	75,872	1,451,830	94.77%
Patient 24-3rd relapse	808,265	63,408	871,673	92.73%
Patient 24-Transformation	1,085,563	71,840	1,157,403	93.79%
Patient 23-Transformation	1,220,139	79,675	1,299,814	93.87%

Table 1.3: *Summary of merged reads for each patient sample.*

Sample	PF Clusters	Amplicon Size	Estimated Coverage
Patient 24- 1st relapse	1,320,495	300	2,200,825
Patient 23- Diagnosis	1,319,287	300	2,198,812
Patient 25-2nd relapse	1,188,214	300	1,980,357
Patient 25-3rd relapse	944,471	300	1,574,118
Patient 25- Transformation	1,546,380	300	2,577,300
Patient 24-3rd relapse	929,244	300	1,548,740
Patient 24- Transformation	1,224,512	300	2,040,853
Patient 23- Transformation	1,365,431	300	2,275,718

Table 1.4: *Estimated coverage for each sample.* Values are based on the assumption of an amplicon size of 300bp.

Patient	Disease	% of tumour related counts with all N-gly motif sites present	% of tumour related counts without all N-gly motif sites present
23	Diagnosis	99.32	0.68
	Transformation	99.42	0.58
24	1st relapse	94.7	5.3*
	3rd relapse	94.8	5.2*
	Transformation	86.7	13.3*
25	2nd relapse	99.15	0.85
	3rd relapse	99.01	1
	Transformation	99.3	0.7

Table 1.5: *Total count numbers for disease event for Patients 23-25.* High % of counts containing motifs reflect how they are a significant feature of the tumour bulk. * For Patient 24, the values take into account only clones where all four sites are present or clones where ≥1 motif are absent. Patient 26, 27 and 28 were excluded due to lack of available information regarding count numbers.

Shared negative clone	% of total tumour count in FL2	% of total tumour count in FL3
1	0.042152	0.037872
2	0.040614	0.056808
3	0.034152	0.067460
4	0.034152	0.035505
5	0.031383	0.031955
6	0.030153	0.056808
7	0.027999	0.018936
8	0.027383	0.027221
9	0.019999	0.036689
10	0.016615	0.024854
11	0.013846	0.015386
12	0.010153	0.021303
13	0.008307	0.040239
14	0.008307	0.026037
15	0.007384	0.020120
16	0.007077	0.014202
17	0.006461	0.014202
18	0.006461	0.023670
19	0.006154	0.023670
20	0.004923	0.017753

Table 1.6: *% of total tumour count for the negative shared subclones of Patient 25.*

		CD21 GC	CD21 Tissue	CD68+ GC	CD68+ Tissue	CD163+ GC	CD163+ Tissue	CD68+ CD163+GC	CD68+CD163+Tissue	CD68-CD163-CD21- GC	CD68-CD163-CD21- Tissue
Reactive LN	Patient 1	3022	584	10	77	201	4664	24	2116	888	4262
	Patient 2	14729	1392	167	2277	529	7853	242	10316	2477	18217
	Patient 3	3778	8078	29	56	211	17868	4	993	340	21181
	Patient 4	25335	9852	84	119	135	2722	7	32	1842	17937
	Patient 5	37371	709	34	172	709	2358	10	28	3026	31112
	Patient 6	7194	718	31	51	1089	6589	46	214	6033	15314
	Patient 7	10204	1640	239	496	308	11741	141	1911	2096	16990
	Patient 8	26656	1069	109	473	1820	6582	89	2804	4605	15937
	Patient 9	3919	202	0	0	294	1251	13	99	215	820
	Total	132208	24244	703	3721	5296	61628	576	18513	21522	141770
5 yr	Patient 1	1489	2201	202	2148	0	201	11	426	132	5052
	Patient 2	4153	895	131	617	137	1765	4	193	133	1719
	Patient 3	15802	3082	1456	1361	3327	3122	1101	1523	746	1060
	Patient 4	37824	388	1134	582	516	7619	171	1563	240	478
	Patient 5	22863	3088	26	60	3708	22541	84	425	1461	9448
	Patient 6	2507	1284	2	4	2320	1514	0	0	1064	751
	Patient 7	33120	2273	4	38	82	305	1	1	369	2591
	Total	117758	13211	2955	4810	10090	37067	1372	4131	4145	21099
15 yr	Patient 1	18338	5463	510	385	4071	7610	82	368	3294	7747
	Patient 2	3094	933	84	577	124	2647	71	660	594	7322
	Patient 3	3426	72	465	506	161	3689	8	179	1368	4654
	Patient 4	65607	159	4736	921	9839	15652	1569	2014	7496	2586
	Patient 5	28	0	1165	333	10769	7169	115	29	1321	1687
	Patient 6	22767	425	380	281	2452	11046	2	13	1296	9291
	Patient 7	10239	1754	93	672	146	396	0	110	953	4518
	Patient 8	3132	73	10	1	402	3412	0	0	674	6869
	Patient 9	10690	1026	89	58	1223	4051	7	1	1019	3896
	Total	137321	9905	7532	3734	29187	55672	1854	3374	18015	48570

Table 1.7: *Count numbers for each cell type identified by the Visiopharm software.*

Supplementary Figures

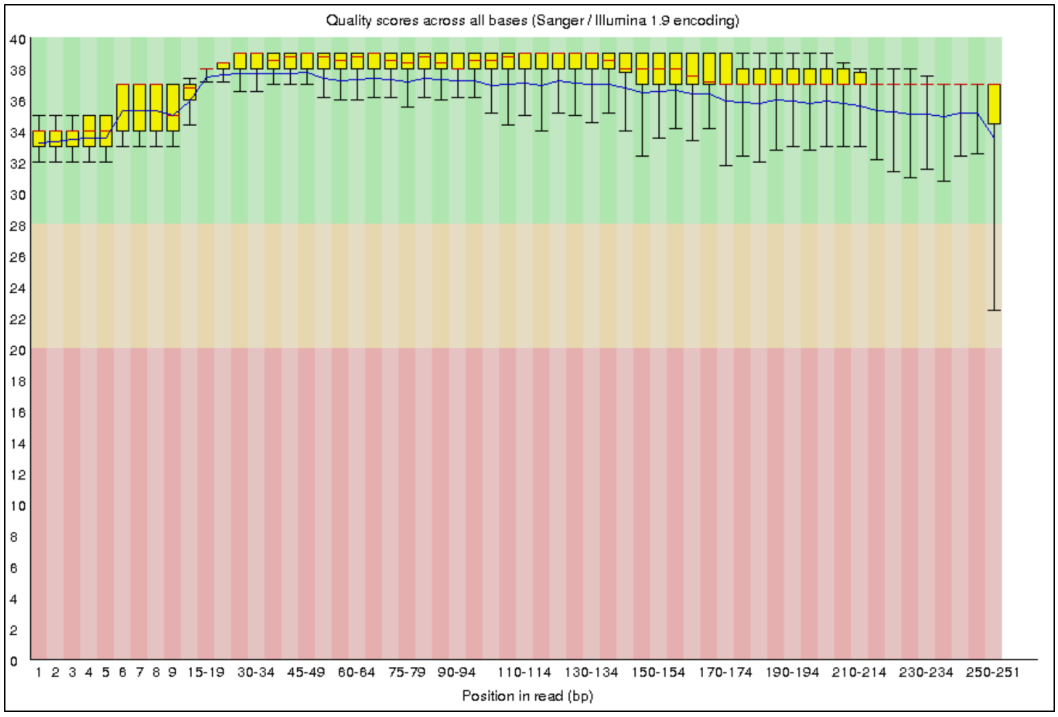


Figure 1.1: *Q-Score distribution of NGS samples (Patients 23-25) after trimming.* X-axis, the position of sequence. Y-axis, average Q scores.

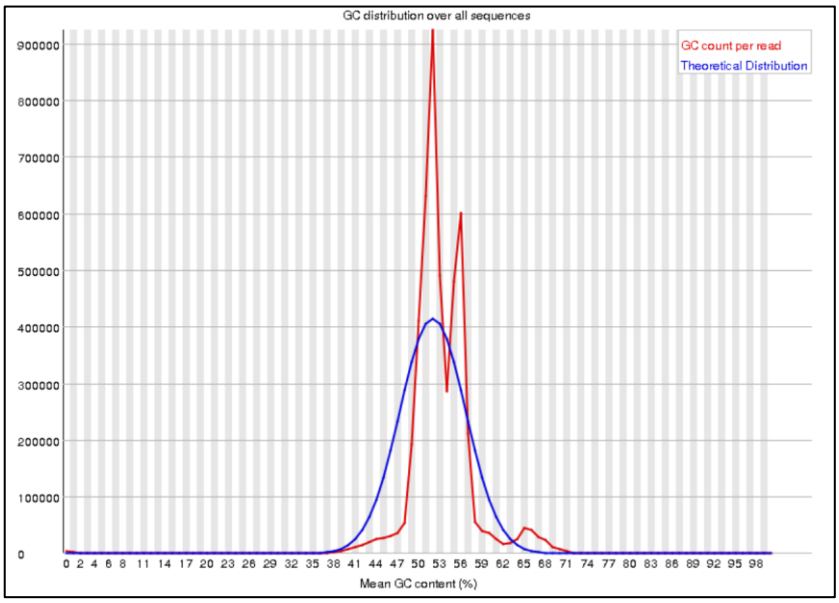


Figure 1.2: *Distribution of GC-content in all samples after trimming.* X-axis, relative GC-content of a sequence in percentage. Y-axis, number of sequences featuring particular GC-percentages normalized to the total number of sequences.

Bibliography

AARTS, W. M., BENDE, R. J., BOSSENBROEK, J. G., PALS, S. T. & VAN NOESEL, C. J. 2001. Variable heavy-chain gene analysis of follicular lymphomas: subclone selection rather than clonal evolution over time. *Blood,* 98**,** 238-40.

AARTS, W. M., BENDE, R. J., STEENBERGEN, E. J., KLUIN, P. M., OOMS, E. C., PALS, S. T. & VAN NOESEL, C. J. 2000. Variable heavy chain gene analysis of follicular lymphomas: correlation between heavy chain isotype expression and somatic mutation load. *Blood,* 95**,** 2922-9.

ADAM, P., KATZENBERGER, T., EIFERT, M., OTT, M. M., ROSENWALD, A., MULLER-HERMELINK, H. K. & OTT, G. 2005. Presence of preserved reactive germinal centers in follicular lymphoma is a strong histopathologic indicator of limited disease stage. *Am J Surg Pathol,* 29**,** 1661-4.

AEBI, M. 2013. N-linked protein glycosylation in the ER. *Biochim Biophys Acta,* 1833**,** 2430-7.

AGOPIAN, J., NAVARRO, J. M., GAC, A. C., LECLUSE, Y., BRIAND, M., GRENOT, P., GAUDUCHON, P., RUMINY, P., LEBAILLY, P., NADEL, B. & ROULLAND, S. 2009. Agricultural pesticide exposure and the molecular connection to lymphomagenesis. *J Exp Med,* 206**,** 1473-83.

AKASAKA, T., AKASAKA, H., YONETANI, N., OHNO, H., YAMABE, H., FUKUHARA, S. & OKUMA, M. 1998. Refinement of the BCL2/immunoglobulin heavy chain fusion gene in t(14;18)(q32;q21) by polymerase chain reaction amplification for long targets. *Genes Chromosomes Cancer,* 21**,** 17-29.

AKASAKA, T., LOSSOS, I. S. & LEVY, R. 2003. BCL6 gene translocation in follicular lymphoma: a harbinger of eventual transformation to diffuse aggressive lymphoma. *Blood,* 102**,** 1443-8.

AL-TOURAH, A. J., GILL, K. K., CHHANABHAI, M., HOSKINS, P. J., KLASA, R. J., SAVAGE, K. J., SEHN, L. H., SHENKIER, T. N., GASCOYNE, R. D. & CONNORS, J. M. 2008. Population-based analysis of incidence and outcome of transformed non-Hodgkin's lymphoma. *J Clin Oncol,* 26**,** 5165-9.

ALAMYAR, E., DUROUX, P., LEFRANC, M. P. & GIUDICELLI, V. 2012. IMGT((R)) tools for the nucleotide analysis of immunoglobulin (IG) and T cell receptor (TR) V-(D)-J repertoires, polymorphisms, and IG mutations: IMGT/V-QUEST and IMGT/HighV-QUEST for NGS. *Methods Mol Biol,* 882**,** 569-604.

ALCOCEBA, M., SEBASTIÁN, E., BALANZATEGUI, A., MARÍN, L., MONTES-MORENO, S., FLORES, T., PUIG, N., SARASQUETE, M. E., CHILLÓN, M. C., JIMÉNEZ, C., CORRAL, R., GONZÁLEZ-BARCA, E., CABALLERO, M. D., MIGUEL, J. F. S., GARCÍA-SANZ, R. & GONZÁLEZ, M. 2012. Preferential Acquision of N-Glycosylation Sites in the VDJ Region in Germinal Center B-Cell-Like Difusse Large B-Cell Lymphoma. *Blood,* 120**,** 1589-1589.

ALIMZHANOV, M. B., KUPRASH, D. V., KOSCO-VILBOIS, M. H., LUZ, A., TURETSKAYA, R. L., TARAKHOVSKY, A., RAJEWSKY, K., NEDOSPASOV, S. A. & PFEFFER, K. 1997. Abnormal development of secondary lymphoid tissues in lymphotoxin beta-deficient mice. *Proc Natl Acad Sci U S A,* 94**,** 9302-7.

ALLAVENA, P., CHIEPPA, M., BIANCHI, G., SOLINAS, G., FABBRI, M., LASKARIN, G. & MANTOVANI, A. 2010. Engagement of the mannose receptor by tumoral mucins activates an immune suppressive phenotype in human tumor-associated macrophages. *Clin Dev Immunol,* 2010**,** 547179.

ALLEN, C. D. & CYSTER, J. G. 2008. Follicular dendritic cell networks of primary follicles and germinal centers: phenotype and function. *Semin Immunol,* 20**,** 14-25.

ALONSO-ALVAREZ, S., MAGNANO, L., ALCOCEBA, M., ANDRADE-CAMPOS, M., ESPINOSA-LARA, N., RODRIGUEZ, G., MERCADAL, S., CARRO, I., SANCHO, J. M., MORENO, M., SALAR, A., GARCIA-PALLAROLS, F., ARRANZ, R., CANNATA, J., TEROL, M. J., TERUEL, A. I., RODRIGUEZ, A., JIMENEZ-UBIETO, A., GONZALEZ DE VILLAMBROSIA, S., BELLO, J. L., LOPEZ, L., MONSALVO, S., NOVELLI, S., DE CABO, E., INFANTE, M. S., PARDAL, E., GARCIA-ALVAREZ, M., DELGADO, J., GONZALEZ, M., MARTIN, A., LOPEZ-GUILLERMO, A. & CABALLERO, M. D. 2017. Risk of, and survival following, histological transformation in follicular lymphoma in the rituximab era. A retrospective multicentre study by the Spanish GELTAMO group. *Br J Haematol,* 178**,** 699-708.

AME-THOMAS, P., MABY-EL HAJJAMI, H., MONVOISIN, C., JEAN, R., MONNIER, D., CAULET-MAUGENDRE, S., GUILLAUDEUX, T., LAMY, T., FEST, T. & TARTE, K. 2007. Human mesenchymal stem cells isolated from bone marrow and lymphoid organs support tumor B-cell growth: role of stromal cells in follicular lymphoma pathogenesis. *Blood,* 109**,** 693-702.

AME-THOMAS, P. & TARTE, K. 2014. The yin and the yang of follicular lymphoma cell niches: role of microenvironment heterogeneity and plasticity. *Semin Cancer Biol,* 24**,** 23-32.

AMIN, R., MOURCIN, F., UHEL, F., PANGAULT, C., RUMINY, P., DUPRE, L., GUIRRIEC, M., MARCHAND, T., FEST, T., LAMY, T. & TARTE, K. 2015. DC-SIGN-expressing macrophages trigger activation of mannosylated IgM B-cell receptor in follicular lymphoma. *Blood,* 126**,** 1911-20.

ARAF, S. & FITZGIBBON, J. 2016. Pediatric-type FL: simply different. *Blood,* 128**,** 1030-1.

ARAF, S., OKOSUN, J., KONIALI, L., FITZGIBBON, J. & HEWARD, J. 2016. Epigenetic dysregulation in follicular lymphoma. *Epigenomics,* 8**,** 77-84.

ARAF, S., WANG, J., KORFI, K., PANGAULT, C., KOTSIOU, E., RIO-MACHIN, A., RAHIM, T., HEWARD, J., CLEAR, A., IQBAL, S., DAVIES, J. K., JOHNSON, P., CALAMINICI, M., MONTOTO, S., AUER, R., CHELALA, C., GRIBBEN, J. G., GRAHAM, T. A., FEST, T., FITZGIBBON, J. & OKOSUN, J. 2018. Genomic profiling reveals spatial intra-tumor heterogeneity in follicular lymphoma. *Leukemia,* 32**,** 1258-1263.

ARDESHNA, K. M., QIAN, W., SMITH, P., BRAGANCA, N., LOWRY, L., PATRICK, P., WARDEN, J., STEVENS, L., POCOCK, C. F., MIALL, F., CUNNINGHAM, D., DAVIES, J., JACK, A., STEPHENS, R., WALEWSKI, J., FERHANOGLU, B., BRADSTOCK, K. & LINCH, D. C. 2014. Rituximab versus a watch-and-wait approach in patients with advanced-stage, asymptomatic, non-bulky follicular lymphoma: an open-label randomised phase 3 trial. *Lancet Oncol,* 15**,** 424-35.

ASTER, J. C. & LONGTINE, J. A. 2002. Detection of BCL2 rearrangements in follicular lymphoma. *Am J Pathol,* 160**,** 759-63.

BACCARELLI, A., HIRT, C., PESATORI, A. C., CONSONNI, D., PATTERSON, D. G., JR., BERTAZZI, P. A., DOLKEN, G. & LANDI, M. T. 2006. t(14;18) translocations in lymphocytes of healthy dioxin-exposed individuals from Seveso, Italy. *Carcinogenesis,* 27**,** 2001-7.

BAHLER, D. W., CAMPBELL, M. J., HART, S., MILLER, R. A., LEVY, S. & LEVY, R. 1991. Ig VH gene expression among human follicular lymphomas. *Blood,* 78**,** 1561-8.

BALL, J. A., CLEAR, A. J., CALAMINICI, M., STAGG, A., LINDSAY, J., GRIBBEN, J. G. & DAVIES, J. 2017. Human Gastro-Intestinal Graft-Versus-Host Disease Is Mediated By Retinoic Acid-Responsive CD8⁺ Effector T-Cells Under IL-23 Polarising Conditions. *Blood,* 130**,** 77-77.

BAN-HOEFEN, M., VANDERPLAS, A., CROSBY-THOMPSON, A. L., ABEL, G. A., CZUCZMAN, M. S., GORDON, L. I., KAMINSKI, M. S., KELLY, J., MILLENSON, M., NADEMANEE, A. P., RODRIGUEZ, M. A., ZELENETZ, A. D., NILAND, J., LACASCE, A. S. & FRIEDBERG,

J. W. 2013. Transformed non-Hodgkin lymphoma in the rituximab era: analysis of the NCCN outcomes database. *Br J Haematol,* 163**,** 487-95.

BARAK, M., ZUCKERMAN, N. S., EDELMAN, H., UNGER, R. & MEHR, R. 2008. IgTree: creating Immunoglobulin variable region gene lineage trees. *J Immunol Methods,* 338**,** 67-74.

BASSO, K. & DALLA-FAVERA, R. 2012. Roles of BCL6 in normal and transformed germinal center B cells. *Immunol Rev,* 247**,** 172-83.

BASTARD, C., DEWEINDT, C., KERCKAERT, J. P., LENORMAND, B., ROSSI, A., PEZZELLA, F., FRUCHART, C., DUVAL, C., MONCONDUIT, M. & TILLY, H. 1994. LAZ3 rearrangements in non-Hodgkin's lymphoma: correlation with histology, immunophenotype, karyotype, and clinical outcome in 217 patients. *Blood,* 83**,** 2423-7.

BASTION, Y., SEBBAN, C., BERGER, F., FELMAN, P., SALLES, G., DUMONTET, C., BRYON, P. A. & COIFFIER, B. 1997. Incidence, predictive factors, and outcome of lymphoma transformation in follicular lymphoma patients. *J Clin Oncol,* 15**,** 1587-94.

BEGUELIN, W., POPOVIC, R., TEATER, M., JIANG, Y., BUNTING, K. L., ROSEN, M., SHEN, H., YANG, S. N., WANG, L., EZPONDA, T., MARTINEZ-GARCIA, E., ZHANG, H., ZHENG, Y., VERMA, S. K., MCCABE, M. T., OTT, H. M., VAN ALLER, G. S., KRUGER, R. G., LIU, Y., MCHUGH, C. F., SCOTT, D. W., CHUNG, Y. R., KELLEHER, N., SHAKNOVICH, R., CREASY, C. L., GASCOYNE, R. D., WONG, K. K., CERCHIETTI, L., LEVINE, R. L., ABDEL-WAHAB, O., LICHT, J. D., ELEMENTO, O. & MELNICK, A. M. 2013. EZH2 is required for germinal center formation and somatic EZH2 mutations promote lymphoid transformation. *Cancer Cell,* 23**,** 677-92.

BENEDETTI, F., BOCCOMINI, C., PATTI, C., BARBUI, A. M., PULSONI, A., MUSSO, M., LIBERATI, A. M., GINI, G., CASTELLINO, C., ROSSINI, F., CICERI, F., SCALABRINI, D. R., STELITANO, C., DI RAIMONDO, F., PERRONE, T., TUCCI, A., BILLIO, A., ZALLIO, F., ZOLI, V., CONGIU, A., NARNI, F., DONDI, A., PARVIS, G., SEMENZATO, G., CORRADINI, P., BRUNA, R., GUELI, A., MANTOAN, B., PASSERA, R., MAGNI, M. & LADETTO, M. 2013. Prolonged Survival Of Poor Risk Follicular Lymphoma Patients Following Primary Treatment With Rituximab-Supplemented CHOP Or HDS With Autograft: Long-Term Results Of The Multicenter Randomized GITMO/FIL Trial. *Blood,* 122**,** 551-551.

BERG, T., THOENE, S., YAP, D., WEE, T., SCHOELER, N., ROSTEN, P., LIM, E., BILENKY, M., MUNGALL, A. J., OELLERICH, T., LEE, S., LAI, C. K., UMLANDT, P., SALMI, A., CHANG, H., YUE, L., LAI, D., CHENG, S. W., MORIN, R. D., HIRST, M., SERVE, H., MARRA, M. A., MORIN, G. B., GASCOYNE, R. D., APARICIO, S. A. & HUMPHRIES, R. K. 2014. A transgenic mouse model demonstrating the oncogenic role of mutations in the polycomb-group gene EZH2 in lymphomagenesis. *Blood,* 123**,** 3914-24.

BERGET, E., MOLVEN, A., LOKELAND, T., HELGELAND, L. & VINTERMYR, O. K. 2015. IGHV gene usage and mutational status in follicular lymphoma: Correlations with prognosis and patient age. *Leuk Res,* 39**,** 702-8.

BOGNAR, A., CSERNUS, B., BODOR, C., REINIGER, L., SZEPESI, A., TOTH, E., KOPPER, L. & MATOLCSY, A. 2005. Clonal selection in the bone marrow involvement of follicular lymphoma. *Leukemia,* 19**,** 1656-62.

BOICE, M., SALLOUM, D., MOURCIN, F., SANGHVI, V., AMIN, R., ORICCHIO, E., JIANG, M., MOTTOK, A., DENIS-LAGACHE, N., CIRIELLO, G., TAM, W., TERUYA-FELDSTEIN, J., DE STANCHINA, E., CHAN, W. C., MALEK, S. N., ENNISHI, D., BRENTJENS, R. J., GASCOYNE, R. D., COGNE, M., TARTE, K. & WENDEL, H. G. 2016. Loss of the HVEM Tumor Suppressor in Lymphoma and Restoration by Modified CAR-T Cells. *Cell,* 167**,** 405-418.e13.

BOSGA-BOUWER, A. G., VAN IMHOFF, G. W., BOONSTRA, R., VAN DER VEEN, A., HARALAMBIEVA, E., VAN DEN BERG, A., DE JONG, B., KRAUSE, V., PALMER, M. C., COUPLAND, R., KLUIN, P. M., VAN DEN BERG, E. & POPPEMA, S. 2003. Follicular lymphoma grade 3B includes 3 cytogenetically defined subgroups with primary t(14;18), 3q27, or other translocations: t(14;18) and 3q27 are mutually exclusive. *Blood,* 101**,** 1149-54.

BOUSKA, A., MCKEITHAN, T. W., DEFFENBACHER, K. E., LACHEL, C., WRIGHT, G. W., IQBAL, J., SMITH, L. M., ZHANG, W., KUCUK, C., RINALDI, A., BERTONI, F., FITZGIBBON, J., FU, K., WEISENBURGER, D. D., GREINER, T. C., DAVE, B. J., GASCOYNE, R. D., ROSENWALD, A., OTT, G., CAMPO, E., RIMSZA, L. M., DELABIE, J., JAFFE, E. S., BRAZIEL, R. M., CONNORS, J. M., STAUDT, L. M. & CHAN, W. C. 2014. Genome-wide copy-number analyses reveal genomic abnormalities involved in transformation of follicular lymphoma. *Blood,* 123**,** 1681-90.

BOYD, S. D. & JOSHI, S. A. 2014. High-Throughput DNA Sequencing Analysis of Antibody Repertoires. *Microbiol Spectr,* 2.

BRAUNINGER, A., HANSMANN, M. L., STRICKLER, J. G., DUMMER, R., BURG, G., RAJEWSKY, K. & KUPPERS, R. 1999. Identification of common germinal-center B-cell precursors in two patients with both Hodgkin's disease and non-Hodgkin's lymphoma. *N Engl J Med,* 340**,** 1239-47.

BREZINSCHEK, H. P., FOSTER, S. J., BREZINSCHEK, R. I., DORNER, T., DOMIATI-SAAD, R. & LIPSKY, P. E. 1997. Analysis of the human VH gene repertoire. Differential effects of selection and somatic hypermutation on human peripheral CD5(+)/IgM+ and CD5(-)/IgM+ B cells. *J Clin Invest,* 99**,** 2488-501.

BRICE, P., BASTION, Y., LEPAGE, E., BROUSSE, N., HAIOUN, C., MOREAU, P., STRAETMANS, N., TILLY, H., TABAH, I. & SOLAL-CELIGNY, P. 1997. Comparison in low-tumor-burden follicular lymphomas between an initial no-treatment policy, prednimustine, or interferon alfa: a randomized study from the Groupe d'Etude des Lymphomes Folliculaires. Groupe d'Etude des Lymphomes de l'Adulte. *J Clin Oncol,* 15**,** 1110-7.

BUECHLER, C., RITTER, M., ORSO, E., LANGMANN, T., KLUCKEN, J. & SCHMITZ, G. 2000. Regulation of scavenger receptor CD163 expression in human monocytes and macrophages by pro- and antiinflammatory stimuli. *J Leukoc Biol,* 67**,** 97-103.

BYERS, R. J., SAKHINIA, E., JOSEPH, P., GLENNIE, C., HOYLAND, J. A., MENASCE, L. P., RADFORD, J. A. & ILLIDGE, T. 2008. Clinical quantitation of immune signature in follicular lymphoma by RT-PCR-based gene expression profiling. *Blood,* 111**,** 4764-70.

CAMPO, E., SWERDLOW, S. H., HARRIS, N. L., PILERI, S., STEIN, H. & JAFFE, E. S. 2011. The 2008 WHO classification of lymphoid neoplasms and beyond: evolving concepts and practical applications. *Blood,* 117**,** 5019-32.

CARBONE, A., GLOGHINI, A., CABRAS, A. & ELIA, G. 2009. The Germinal centre-derived lymphomas seen through their cellular microenvironment. *Br J Haematol,* 145**,** 468-80.

CARIAPPA, A., TANG, M., PARNG, C., NEBELITSKIY, E., CARROLL, M., GEORGOPOULOS, K. & PILLAI, S. 2001. The follicular versus marginal zone B lymphocyte cell fate decision is regulated by Aiolos, Btk, and CD21. *Immunity,* 14**,** 603-15.

CARLOTTI, E., WRENCH, D., MATTHEWS, J., IQBAL, S., DAVIES, A., NORTON, A., HART, J., LAI, R., MONTOTO, S., GRIBBEN, J. G., LISTER, T. A. & FITZGIBBON, J. 2009. Transformation of follicular lymphoma to diffuse large B-cell lymphoma may occur by divergent evolution from a common progenitor cell or by direct evolution from the follicular lymphoma clone. *Blood,* 113**,** 3553-7.

CARLOTTI, E., WRENCH, D., ROSIGNOLI, G., MARZEC, J., SANGARALINGAM, A., HAZANOV, L., MICHAELI, M., HALLAM, S., CHAPLIN, T., IQBAL, S., CALAMINICI, M., YOUNG, B., MEHR, R., CAMPBELL, P., FITZGIBBON, J. & GRIBBEN, J. G. 2015. High Throughput Sequencing Analysis of the Immunoglobulin Heavy Chain Gene from Flow-Sorted B Cell Sub-Populations Define the Dynamics of Follicular Lymphoma Clonal Evolution. *PLoS One,* 10**,** e0134833.

CARRERAS, J., LOPEZ-GUILLERMO, A., RONCADOR, G., VILLAMOR, N., COLOMO, L., MARTINEZ, A., HAMOUDI, R., HOWAT, W. J., MONTSERRAT, E. & CAMPO, E. 2009. High numbers of tumor-infiltrating programmed cell death 1-positive regulatory lymphocytes are associated with improved overall survival in follicular lymphoma. *J Clin Oncol,* 27**,** 1470-6.

CASULO, C. 2016. Prognostic factors in follicular lymphoma: new tools to personalize risk. *Hematology Am Soc Hematol Educ Program,* 2016**,** 269-276.

CASULO, C., BYRTEK, M., DAWSON, K. L., ZHOU, X., FARBER, C. M., FLOWERS, C. R., HAINSWORTH, J. D., MAURER, M. J., CERHAN, J. R., LINK, B. K., ZELENETZ, A. D. & FRIEDBERG, J. W. 2015. Early Relapse of Follicular Lymphoma After Rituximab Plus Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone Defines Patients at High Risk for Death: An Analysis From the National LymphoCare Study. *J Clin Oncol,* 33**,** 2516-22.

CHA, S. C., QIN, H., KANNAN, S., RAWAL, S., WATKINS, L. S., BAIO, F. E., WU, W., ONG, J., WEI, J., KWAK, B., KIM, S., POPESCU, M. S., PAICK, D. S., KIM, K., LUONG, A., DAVIS, R. E., SCHROEDER, H. W., JR., KWAK, L. W. & NEELAPU, S. S. 2013. Nonstereotyped lymphoma B cell receptors recognize vimentin as a shared autoantigen. *J Immunol,* 190**,** 4887-98.

CHEUNG, K. J., DELANEY, A., BEN-NERIAH, S., SCHEIN, J., LEE, T., SHAH, S. P., CHEUNG, D., JOHNSON, N. A., MUNGALL, A. J., TELENIUS, A., LAI, B., BOYLE, M., CONNORS, J. M., GASCOYNE, R. D., MARRA, M. A. & HORSMAN, D. E. 2010a. High resolution analysis of follicular lymphoma genomes reveals somatic recurrent sites of copy-neutral loss of heterozygosity and copy number alterations that target single genes. *Genes Chromosomes Cancer,* 49**,** 669-81.

CHEUNG, K. J., JOHNSON, N. A., AFFLECK, J. G., SEVERSON, T., STEIDL, C., BEN-NERIAH, S., SCHEIN, J., MORIN, R. D., MOORE, R., SHAH, S. P., QIAN, H., PAUL, J. E., TELENIUS, A., RELANDER, T., LAM, W., SAVAGE, K., CONNORS, J. M., BROWN, C., MARRA, M. A., GASCOYNE, R. D. & HORSMAN, D. E. 2010b. Acquired TNFRSF14 mutations in follicular lymphoma are associated with worse prognosis. *Cancer Res,* 70**,** 9166-74.

CHEUNG, K. J., SHAH, S. P., STEIDL, C., JOHNSON, N., RELANDER, T., TELENIUS, A., LAI, B., MURPHY, K. P., LAM, W., AL-TOURAH, A. J., CONNORS, J. M., NG, R. T., GASCOYNE, R. D. & HORSMAN, D. E. 2009a. Genome-wide profiling of follicular lymphoma by array comparative genomic hybridization reveals prognostically significant DNA copy number imbalances. *Blood,* 113**,** 137-48.

CHEUNG, M. C., BAILEY, D., PENNELL, N., IMRIE, K. R., BERINSTEIN, N. L., AMATO, D. & GHORAB, Z. 2009b. In situ localization of follicular lymphoma: evidence for subclinical systemic disease with detection of an identical BCL-2/IGH fusion gene in blood and lymph node. *Leukemia,* 23**,** 1176-9.

CHIEPPA, M., BIANCHI, G., DONI, A., DEL PRETE, A., SIRONI, M., LASKARIN, G., MONTI, P., PIEMONTI, L., BIONDI, A., MANTOVANI, A., INTRONA, M. & ALLAVENA, P. 2003. Cross-linking of the mannose receptor on monocyte-derived dendritic cells activates an anti-inflammatory immunosuppressive program. *J Immunol,* 171**,** 4552-60.

CHIU, B. C., DAVE, B. J., BLAIR, A., GAPSTUR, S. M., ZAHM, S. H. & WEISENBURGER, D. D. 2006. Agricultural pesticide use and risk of t(14;18)-defined subtypes of non-Hodgkin lymphoma. *Blood,* 108**,** 1363-9.

CHOUDHURY, K. R., YAGLE, K. J., SWANSON, P. E., KROHN, K. A. & RAJENDRAN, J. G. 2010. A robust automated measure of average antibody staining in immunohistochemistry images. *J Histochem Cytochem,* 58**,** 95-107.

CLEAR, A. J., LEE, A. M., CALAMINICI, M., RAMSAY, A. G., MORRIS, K. J., HALLAM, S., KELLY, G., MACDOUGALL, F., LISTER, T. A. & GRIBBEN, J. G. 2010. Increased angiogenic sprouting in poor prognosis FL is associated with elevated numbers of CD163+ macrophages within the immediate sprouting microenvironment. *Blood,* 115**,** 5053-6.

CLEARY, M. L., GALILI, N. & SKLAR, J. 1986a. Detection of a second t(14;18) breakpoint cluster region in human follicular lymphomas. *J Exp Med,* 164**,** 315-20.

CLEARY, M. L., MEEKER, T. C., LEVY, S., LEE, E., TRELA, M., SKLAR, J. & LEVY, R. 1986b. Clustering of extensive somatic mutations in the variable region of an immunoglobulin heavy chain gene from a human B cell lymphoma. *Cell,* 44**,** 97-106.

CLEARY, M. L. & SKLAR, J. 1985. Nucleotide sequence of a t(14;18) chromosomal breakpoint in follicular lymphoma and demonstration of a breakpoint-cluster region near a transcriptionally active locus on chromosome 18. *Proc Natl Acad Sci U S A,* 82**,** 7439-43.

COELHO, V., KRYSOV, S., GHAEMMAGHAMI, A. M., EMARA, M., POTTER, K. N., JOHNSON, P., PACKHAM, G., MARTINEZ-POMARES, L. & STEVENSON, F. K. 2010. Glycosylation of surface Ig creates a functional bridge between human follicular lymphoma and microenvironmental lectins. *Proc Natl Acad Sci U S A,* 107**,** 18587-92.

COMPAGNO, M., LIM, W. K., GRUNN, A., NANDULA, S. V., BRAHMACHARY, M., SHEN, Q., BERTONI, F., PONZONI, M., SCANDURRA, M., CALIFANO, A., BHAGAT, G., CHADBURN, A., DALLA-FAVERA, R. & PASQUALUCCI, L. 2009. Mutations of multiple genes cause deregulation of NF-kappaB in diffuse large B-cell lymphoma. *Nature,* 459**,** 717-21.

CONCONI, A., PONZIO, C., LOBETTI-BODONI, C., MOTTA, M., RANCOITA, P. M., STATHIS, A., MOCCIA, A. A., MAZZUCCHELLI, L., BERTONI, F., GHIELMINI, M., CAVALLI, F. & ZUCCA, E. 2012. Incidence, risk factors and outcome of histological transformation in follicular lymphoma. *Br J Haematol,* 157**,** 188-96.

CONG, P., RAFFELD, M., TERUYA-FELDSTEIN, J., SORBARA, L., PITTALUGA, S. & JAFFE, E. S. 2002. In situ localization of follicular lymphoma: description and analysis by laser capture microdissection. *Blood,* 99**,** 3376-82.

COSSMAN, J., NECKERS, L. M., HSU, S., LONGO, D. & JAFFE, E. S. 1984. Low-grade lymphomas. Expression of developmentally regulated B-cell antigens. *Am J Pathol,* 115**,** 117-24.

COSTELLO, R. T., MALLET, F., BARBARAT, B., SCHIANO DE COLELLA, J. M., SAINTY, D., SWEET, R. W., TRUNEH, A. & OLIVE, D. 2003. Stimulation of non-Hodgkin's lymphoma via HVEM: an alternate and safe way to increase Fas-induced apoptosis and improve tumor immunogenicity. *Leukemia,* 17**,** 2500-7.

CROTTY, S. 2014. T follicular helper cell differentiation, function, and roles in disease. *Immunity,* 41**,** 529-42.

CZUCZMAN, M. S., WEAVER, R., ALKUZWENY, B., BERLFEIN, J. & GRILLO-LOPEZ, A. J. 2004. Prolonged clinical and molecular remission in patients with low-grade or follicular non-Hodgkin's lymphoma treated with rituximab plus CHOP chemotherapy: 9-year follow-up. *J Clin Oncol,* 22**,** 4711-6.

DAL PORTO, J. M., GAULD, S. B., MERRELL, K. T., MILLS, D., PUGH-BERNARD, A. E. & CAMBIER, J. 2004. B cell antigen receptor signaling 101. *Mol Immunol,* 41**,** 599-613.

DAMBUZA, I. M. & BROWN, G. D. 2015. C-type lectins in immunity: recent developments. *Curr Opin Immunol,* 32**,** 21-7.

DARRINGTON, D. L., VOSE, J. M., ANDERSON, J. R., BIERMAN, P. J., BISHOP, M. R., CHAN, W. C., MORRIS, M. E., REED, E. C., SANGER, W. G., TARANTOLO, S. R. & ET AL. 1994. Incidence and characterization of secondary myelodysplastic syndrome and acute myelogenous leukemia following high-dose chemoradiotherapy and autologous stem-cell transplantation for lymphoid malignancies. *J Clin Oncol,* 12**,** 2527-34.

DAVE, S. S., WRIGHT, G., TAN, B., ROSENWALD, A., GASCOYNE, R. D., CHAN, W. C., FISHER, R. I., BRAZIEL, R. M., RIMSZA, L. M., GROGAN, T. M., MILLER, T. P., LEBLANC, M., GREINER, T. C., WEISENBURGER, D. D., LYNCH, J. C., VOSE, J., ARMITAGE, J. O., SMELAND, E. B., KVALOY, S., HOLTE, H., DELABIE, J., CONNORS, J. M., LANSDORP, P. M., OUYANG, Q., LISTER, T. A., DAVIES, A. J., NORTON, A. J., MULLER-HERMELINK, H. K., OTT, G., CAMPO, E., MONTSERRAT, E., WILSON, W. H., JAFFE, E. S., SIMON, R., YANG, L., POWELL, J., ZHAO, H., GOLDSCHMIDT, N., CHIORAZZI, M. & STAUDT, L. M. 2004. Prediction of survival in follicular lymphoma based on molecular features of tumor-infiltrating immune cells. *N Engl J Med,* 351**,** 2159-69.

DAVIES, A. J., ROSENWALD, A., WRIGHT, G., LEE, A., LAST, K. W., WEISENBURGER, D. D., CHAN, W. C., DELABIE, J., BRAZIEL, R. M., CAMPO, E., GASCOYNE, R. D., JAFFE, E. S., MULLER-HERMELINK, K., OTT, G., CALAMINICI, M., NORTON, A. J., GOFF, L. K., FITZGIBBON, J., STAUDT, L. M. & ANDREW LISTER, T. 2007. Transformation of follicular lymphoma to diffuse large B-cell lymphoma proceeds by distinct oncogenic mechanisms. *Br J Haematol,* 136**,** 286-93.

DE JONG, D., KOSTER, A., HAGENBEEK, A., RAEMAEKERS, J., VELDHUIZEN, D., HEISTERKAMP, S., DE BOER, J. P. & VAN GLABBEKE, M. 2009. Impact of the tumor microenvironment on prognosis in follicular lymphoma is dependent on specific treatment protocols. *Haematologica,* 94**,** 70-7.

DE LEON, E. D., ALKAN, S., HUANG, J. C. & HSI, E. D. 1998. Usefulness of an immunohistochemical panel in paraffin-embedded tissues for the differentiation of B-cell non-Hodgkin's lymphomas of small lymphocytes. *Mod Pathol,* 11**,** 1046-51.

DE SILVA, N. S. & KLEIN, U. 2015. Dynamics of B cells in germinal centres. *Nat Rev Immunol,* 15**,** 137-48.

DOLKEN, G., ILLERHAUS, G., HIRT, C. & MERTELSMANN, R. 1996. BCL-2/JH rearrangements in circulating B cells of healthy blood donors and patients with nonmalignant diseases. *J Clin Oncol,* 14**,** 1333-44.

EGLE, A., HARRIS, A. W., BATH, M. L., O'REILLY, L. & CORY, S. 2004. VavP-Bcl2 transgenic mice develop follicular lymphoma preceded by germinal center hyperplasia. *Blood,* 103**,** 2276-83.

EIDE, M. B., LIESTOL, K., LINGJAERDE, O. C., HYSTAD, M. E., KRESSE, S. H., MEZA-ZEPEDA, L., MYKLEBOST, O., TROEN, G., AAMOT, H. V., HOLTE, H., SMELAND, E. B. & DELABIE, J. 2010. Genomic alterations reveal potential for higher grade transformation in follicular lymphoma and confirm parallel evolution of tumor cell clones. *Blood,* 116**,** 1489-97.

ENGERING, A. J., CELLA, M., FLUITSMA, D. M., HOEFSMIT, E. C., LANZAVECCHIA, A. & PIETERS, J. 1997. Mannose receptor mediated antigen uptake and presentation in human dendritic cells. *Adv Exp Med Biol,* 417**,** 183-7.

EPELMAN, S., LAVINE, K. J. & RANDOLPH, G. J. 2014. Origin and functions of tissue macrophages. *Immunity,* 41**,** 21-35.

ERLICH, Y., MITRA, P. P., DELABASTIDE, M., MCCOMBIE, W. R. & HANNON, G. J. 2008. Alta-Cyclic: a self-optimizing base caller for next-generation sequencing. *Nat Methods,* 5**,** 679-82.

ESTELLER, M. 2008. Epigenetics in cancer. *N Engl J Med,* 358**,** 1148-59.

EVENS, A. M., BALASUBRAMANIAN, S., VOSE, J. M., HARB, W., GORDON, L. I., LANGDON, R., SPRAGUE, J., SIRISAWAD, M., MANI, C., YUE, J., LUAN, Y., HORTON, S., GRAEF, T. & BARTLETT, N. L. 2016. A Phase I/II Multicenter, Open-Label Study of the Oral Histone Deacetylase Inhibitor Abexinostat in Relapsed/Refractory Lymphoma. *Clin Cancer Res,* 22**,** 1059-66.

FAIS, F., GHIOTTO, F., HASHIMOTO, S., SELLARS, B., VALETTO, A., ALLEN, S. L., SCHULMAN, P., VINCIGUERRA, V. P., RAI, K., RASSENTI, L. Z., KIPPS, T. J., DIGHIERO, G., SCHROEDER, H. W., JR., FERRARINI, M. & CHIORAZZI, N. 1998. Chronic lymphocytic leukemia B cells express restricted sets of mutated and unmutated antigen receptors. *J Clin Invest,* 102**,** 1515-25.

FARINHA, P., MASOUDI, H., SKINNIDER, B. F., SHUMANSKY, K., SPINELLI, J. J., GILL, K., KLASA, R., VOSS, N., CONNORS, J. M. & GASCOYNE, R. D. 2005. Analysis of multiple biomarkers shows that lymphoma-associated macrophage (LAM) content is an independent predictor of survival in follicular lymphoma (FL). *Blood,* 106**,** 2169-74.

FARRELL, K. & JARRETT, R. F. 2011. The molecular pathogenesis of Hodgkin lymphoma. *Histopathology,* 58**,** 15-25.

FEARON, D. T. & CARROLL, M. C. 2000. Regulation of B lymphocyte responses to foreign and self-antigens by the CD19/CD21 complex. *Annu Rev Immunol,* 18**,** 393-422.

FEDERICO, M., BELLEI, M., MARCHESELLI, L., LUMINARI, S., LOPEZ-GUILLERMO, A., VITOLO, U., PRO, B., PILERI, S., PULSONI, A., SOUBEYRAN, P., CORTELAZZO, S., MARTINELLI, G., MARTELLI, M., RIGACCI, L., ARCAINI, L., DI RAIMONDO, F., MERLI, F., SABATTINI, E., MCLAUGHLIN, P. & SOLAL-CELIGNY, P. 2009. Follicular lymphoma international prognostic index 2: a new prognostic index for follicular lymphoma developed by the international follicular lymphoma prognostic factor project. *J Clin Oncol,* 27**,** 4555-62.

FEDERICO, M., LUMINARI, S., DONDI, A., TUCCI, A., VITOLO, U., RIGACCI, L., DI RAIMONDO, F., CARELLA, A. M., PULSONI, A., MERLI, F., ARCAINI, L., ANGRILLI, F., STELITANO, C., GAIDANO, G., DELL'OLIO, M., MARCHESELLI, L., FRANCO, V., GALIMBERTI, S., SACCHI, S. & BRUGIATELLI, M. 2013. R-CVP versus R-CHOP versus R-FM for the initial treatment of patients with advanced-stage follicular lymphoma: results of the FOLL05 trial conducted by the Fondazione Italiana Linfomi. *J Clin Oncol,* 31**,** 1506-13.

FEINBERG, A. P. & VOGELSTEIN, B. 1983. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature,* 301**,** 89-92.

FELSENFELD, G. 2014. A brief history of epigenetics. *Cold Spring Harb Perspect Biol,* 6.

FITZGIBBON, J., IQBAL, S., DAVIES, A., O'SHEA, D., CARLOTTI, E., CHAPLIN, T., MATTHEWS, J., RAGHAVAN, M., NORTON, A., LISTER, T. A. & YOUNG, B. D. 2007. Genome-wide detection of recurring sites of uniparental disomy in follicular and transformed follicular lymphoma. *Leukemia,* 21**,** 1514-20.

FLINN, I. W., VAN DER JAGT, R., KAHL, B. S., WOOD, P., HAWKINS, T. E., MACDONALD, D., HERTZBERG, M., KWAN, Y. L., SIMPSON, D., CRAIG, M., KOLIBABA, K., ISSA, S., CLEMENTI, R., HALLMAN, D. M., MUNTEANU, M., CHEN, L. & BURKE, J. M. 2014. Randomized trial of bendamustine-rituximab or R-CHOP/R-CVP in first-line treatment of indolent NHL or MCL: the BRIGHT study. *Blood,* 123**,** 2944-52.

FORSTPOINTNER, R., DREYLING, M., REPP, R., HERMANN, S., HANEL, A., METZNER, B., POTT, C., HARTMANN, F., ROTHMANN, F., ROHRBERG, R., BOCK, H. P., WANDT, H., UNTERHALT, M. & HIDDEMANN, W. 2004. The addition of rituximab to a combination of fludarabine, cyclophosphamide, mitoxantrone (FCM) significantly increases the response rate and prolongs survival as compared with FCM alone in patients with relapsed and refractory follicular and mantle cell lymphomas: results of a prospective randomized study of the German Low-Grade Lymphoma Study Group. *Blood,* 104**,** 3064-71.

FRIEDBERG, J. W., COHEN, P., CHEN, L., ROBINSON, K. S., FORERO-TORRES, A., LA CASCE, A. S., FAYAD, L. E., BESSUDO, A., CAMACHO, E. S., WILLIAMS, M. E., VAN DER JAGT, R. H., OLIVER, J. W. & CHESON, B. D. 2008. Bendamustine in patients with rituximab-refractory indolent and transformed non-Hodgkin's lymphoma: results from a phase II multicenter, single-agent study. *J Clin Oncol,* 26**,** 204-10.

GALL, E. A. & MALLORY, T. B. 1942. Malignant Lymphoma: A Clinico-Pathologic Survey of 618 Cases. *Am J Pathol,* 18**,** 381-429.

GARCIA-MANERO, G. & FENAUX, P. 2011. Hypomethylating agents and other novel strategies in myelodysplastic syndromes. *J Clin Oncol,* 29**,** 516-23.

GARVIN, A. J., SIMON, R. M., OSBORNE, C. K., MERRILL, J., YOUNG, R. C. & BERARD, C. W. 1983. An autopsy study of histologic progression in non-Hodgkin's lymphomas. 192 cases from the National Cancer Institute. *Cancer,* 52**,** 393-8.

GAVEL, Y. & VON HEIJNE, G. 1990. Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering. *Protein Eng,* 3**,** 433-42.

GAVRIELI, M., WATANABE, N., LOFTIN, S. K., MURPHY, T. L. & MURPHY, K. M. 2003. Characterization of phosphotyrosine binding motifs in the cytoplasmic domain of B and T lymphocyte attenuator required for association with protein tyrosine phosphatases SHP-1 and SHP-2. *Biochem Biophys Res Commun,* 312**,** 1236-43.

GEIJTENBEEK, T. B. & GRINGHUIS, S. I. 2009. Signalling through C-type lectin receptors: shaping immune responses. *Nat Rev Immunol,* 9**,** 465-79.

GEIJTENBEEK, T. B., TORENSMA, R., VAN VLIET, S. J., VAN DUIJNHOVEN, G. C., ADEMA, G. J., VAN KOOYK, Y. & FIGDOR, C. G. 2000. Identification of DC-SIGN, a novel dendritic cell-specific ICAM-3 receptor that supports primary immune responses. *Cell,* 100**,** 575-85.

GHIA, E. M., JAIN, S., WIDHOPF, G. F., 2ND, RASSENTI, L. Z., KEATING, M. J., WIERDA, W. G., GRIBBEN, J. G., BROWN, J. R., RAI, K. R., BYRD, J. C., KAY, N. E., GREAVES, A. W. & KIPPS, T. J. 2008. Use of IGHV3-21 in chronic lymphocytic leukemia is associated with high-risk disease and reflects antigen-driven, post-germinal center leukemogenic selection. *Blood,* 111**,** 5101-8.

GHIA, P., BOUSSIOTIS, V. A., SCHULTZE, J. L., CARDOSO, A. A., DORFMAN, D. M., GRIBBEN, J. G., FREEDMAN, A. S. & NADLER, L. M. 1998. Unbalanced expression of bcl-2 family proteins in follicular lymphoma: contribution of CD40 signaling in promoting survival. *Blood,* 91**,** 244-51.

GLEESON, M., HAWKES, E. A., PECKITT, C., WOTHERSPOON, A., ATTYGALLE, A., SHARMA, B., DU, Y., ETHELL, M., POTTER, M., DEARDEN, C., HORWICH, A., CHAU, I. & CUNNINGHAM, D. 2017. Outcomes for transformed follicular lymphoma in the rituximab era: the Royal Marsden experience 2003-2013. *Leuk Lymphoma,* 58**,** 1805-1813.

GREEN, M. R., GENTLES, A. J., NAIR, R. V., IRISH, J. M., KIHIRA, S., LIU, C. L., KELA, I., HOPMANS, E. S., MYKLEBUST, J. H., JI, H., PLEVRITIS, S. K., LEVY, R. & ALIZADEH, A. A. 2013. Hierarchy in somatic mutations arising during genomic evolution and progression of follicular lymphoma. *Blood,* 121**,** 1604-11.

GREEN, M. R., KIHIRA, S., LIU, C. L., NAIR, R. V., SALARI, R., GENTLES, A. J., IRISH, J., STEHR, H., VICENTE-DUENAS, C., ROMERO-CAMARERO, I., SANCHEZ-GARCIA, I., PLEVRITIS, S. K., ARBER, D. A., BATZOGLOU, S., LEVY, R. & ALIZADEH, A. A. 2015. Mutations in early follicular lymphoma progenitors are associated with suppressed antigen presentation. *Proc Natl Acad Sci U S A,* 112**,** E1116-25.

GRIBBEN, J. G., NEUBERG, D., FREEDMAN, A. S., GIMMI, C. D., PESEK, K. W., BARBER, M., SAPORITO, L., WOO, S. D., CORAL, F., SPECTOR, N. & ET AL. 1993. Detection by polymerase chain reaction of residual cells with the bcl-2 translocation is associated with increased risk of relapse after autologous bone marrow transplantation for B-cell lymphoma. *Blood,* 81**,** 3449-57.

GRILLO, F., BRUZZONE, M., PIGOZZI, S., PROSAPIO, S., MIGLIORA, P., FIOCCA, R. & MASTRACCI, L. 2017. Immunohistochemistry on old archival paraffin blocks: is there an expiry date? *J Clin Pathol,* 70**,** 988-993.

GUADAGNOLO, B. A., LI, S., NEUBERG, D., NG, A., HUA, L., SILVER, B., STEVENSON, M. A. & MAUCH, P. 2006. Long-term outcome and mortality trends in early-stage, Grade 1-2 follicular lymphoma treated with radiation therapy. *Int J Radiat Oncol Biol Phys,* 64**,** 928-34.

GUILLOTON, F., CARON, G., MENARD, C., PANGAULT, C., AME-THOMAS, P., DULONG, J., DE VOS, J., ROSSILLE, D., HENRY, C., LAMY, T., FOUQUET, O., FEST, T. & TARTE, K. 2012. Mesenchymal stromal cells orchestrate follicular lymphoma cell niche through the CCL2-dependent recruitment and polarization of monocytes. *Blood,* 119**,** 2556-67.

GYAN, E., FOUSSARD, C., BERTRAND, P., MICHENET, P., LE GOUILL, S., BERTHOU, C., MAISONNEUVE, H., DELWAIL, V., GRESSIN, R., QUITTET, P., VILQUE, J. P., DESABLENS, B., JAUBERT, J., RAMEE, J. F., ARAKELYAN, N., THYSS, A., MOLUCON-CHABROT, C., DELEPINE, R., MILPIED, N., COLOMBAT, P. & DECONINCK, E. 2009. High-dose therapy followed by autologous purged stem cell transplantation and doxorubicin-based chemotherapy in patients with advanced follicular lymphoma: a randomized multicenter study by the GOELAMS with final results after a median follow-up of 9 years. *Blood,* 113**,** 995-1001.

HARRIS, N. L., NADLER, L. M. & BHAN, A. K. 1984. Immunohistologic characterization of two malignant lymphomas of germinal center type (centroblastic/centrocytic and centrocytic) with monoclonal antibodies. Follicular and diffuse lymphomas of small-cleaved-cell type are related but distinct entities. *Am J Pathol,* 117**,** 262-72.

HARROP, J. A., MCDONNELL, P. C., BRIGHAM-BURKE, M., LYN, S. D., MINTON, J., TAN, K. B., DEDE, K., SPAMPANATO, J., SILVERMAN, C., HENSLEY, P., DIPRINZIO, R., EMERY, J. G., DEEN, K., EICHMAN, C., CHABOT-FLETCHER, M., TRUNEH, A. & YOUNG, P. R. 1998. Herpesvirus entry mediator ligand (HVEM-L), a novel ligand for HVEM/TR2, stimulates proliferation of T cells and inhibits HT29 cell growth. *J Biol Chem,* 273**,** 27548-56.

HASHIMOTO, A., OKADA, H., JIANG, A., KUROSAKI, M., GREENBERG, S., CLARK, E. A. & KUROSAKI, T. 1998. Involvement of guanosine triphosphatases and phospholipase C-gamma2 in extracellular signal-regulated kinase, c-Jun NH2-terminal kinase, and p38 mitogen-activated protein kinase activation by the B cell antigen receptor. *J Exp Med,* 188**,** 1287-95.

HENOPP, T., QUINTANILLA-MARTINEZ, L., FEND, F. & ADAM, P. 2011. Prevalence of follicular lymphoma in situ in consecutively analysed reactive lymph nodes. *Histopathology,* 59**,** 139-42.

HEROLD, M., DOLKEN, G., FIEDLER, F., FRANKE, A., FREUND, M., HELBIG, W. & PASOLD, R. 2003. Randomized phase III study for the treatment of advanced indolent non-

Hodgkin's lymphomas (NHL) and mantle cell lymphoma: chemotherapy versus chemotherapy plus rituximab. *Ann Hematol,* 82**,** 77-9.

HERSHBERG, U., UDUMAN, M., SHLOMCHIK, M. J. & KLEINSTEIN, S. H. 2008. Improved methods for detecting selection by mutation analysis of Ig V region sequences. *Int Immunol,* 20**,** 683-94.

HIDDEMANN, W., KNEBA, M., DREYLING, M., SCHMITZ, N., LENGFELDER, E., SCHMITS, R., REISER, M., METZNER, B., HARDER, H., HEGEWISCH-BECKER, S., FISCHER, T., KROPFF, M., REIS, H. E., FREUND, M., WORMANN, B., FUCHS, R., PLANKER, M., SCHIMKE, J., EIMERMACHER, H., TRUMPER, L., ALDAOUD, A., PARWARESCH, R. & UNTERHALT, M. 2005. Frontline therapy with rituximab added to the combination of cyclophosphamide, doxorubicin, vincristine, and prednisone (CHOP) significantly improves the outcome for patients with advanced-stage follicular lymphoma compared with therapy with CHOP alone: results of a prospective randomized study of the German Low-Grade Lymphoma Study Group. *Blood,* 106**,** 3725-32.

HIRT, C., DOLKEN, G., JANZ, S. & RABKIN, C. S. 2007. Distribution of t(14;18)-positive, putative lymphoma precursor cells among B-cell subsets in healthy individuals. *Br J Haematol,* 138**,** 349-53.

HORN, H., SCHMELTER, C., LEICH, E., SALAVERRIA, I., KATZENBERGER, T., OTT, M. M., KALLA, J., ROMERO, M., SIEBERT, R., ROSENWALD, A. & OTT, G. 2011. Follicular lymphoma grade 3B is a distinct neoplasm according to cytogenetic and immunohistochemical profiles. *Haematologica,* 96**,** 1327-34.

HORN, M., KREUZ, M., KÜPPERS, R. & LOEFFLER, M. 2016. Dynamical Modeling of Clonal Evolution in Paired Primary and Relapsed Follicular Lymphoma Predicts a Link Between Cell Migration and Evolutionary Heterogeneity. *Blood,* 128**,** 2929-2929.

HORNING, S. J. & ROSENBERG, S. A. 1984. The natural history of initially untreated low-grade non-Hodgkin's lymphomas. *N Engl J Med,* 311**,** 1471-5.

HORSMAN, D. E., CONNORS, J. M., PANTZAR, T. & GASCOYNE, R. D. 2001. Analysis of secondary chromosomal alterations in 165 cases of follicular lymphoma with t(14;18). *Genes Chromosomes Cancer,* 30**,** 375-82.

HORSMAN, D. E., OKAMOTO, I., LUDKOVSKI, O., LE, N., HARDER, L., GESK, S., SIEBERT, R., CHHANABHAI, M., SEHN, L., CONNORS, J. M. & GASCOYNE, R. D. 2003. Follicular lymphoma lacking the t(14;18)(q32;q21): identification of two disease subtypes. *Br J Haematol,* 120**,** 424-33.

HUBBARD, S. M., CHABNER, B. A., DEVITA, V. T., JR., SIMON, R., BERARD, C. W., JONES, R. B., GARVIN, A. J., CANELLOS, G. P., OSBORNE, C. K. & YOUNG, R. C. 1982. Histologic progression in non-Hodgkin's lymphoma. *Blood,* 59**,** 258-64.

HUET, S., SUJOBERT, P. & SALLES, G. 2018a. From genetics to the clinic: a translational perspective on follicular lymphoma. *Nat Rev Cancer,* 18**,** 224-239.

HUET, S., TESSON, B., JAIS, J. P., FELDMAN, A. L., MAGNANO, L., THOMAS, E., TRAVERSE-GLEHEN, A., ALBAUD, B., CARRERE, M., XERRI, L., ANSELL, S. M., BASEGGIO, L., REYES, C., TARTE, K., BOYAULT, S., HAIOUN, C., LINK, B. K., FEUGIER, P., LOPEZ-GUILLERMO, A., TILLY, H., BRICE, P., HAYETTE, S., JARDIN, F., OFFNER, F., SUJOBERT, P., GENTIEN, D., VIARI, A., CAMPO, E., CERHAN, J. R. & SALLES, G. 2018b. A gene-expression profiling score for prediction of outcome in patients with follicular lymphoma: a retrospective training and validation analysis in three international cohorts. *Lancet Oncol,* 19**,** 549-561.

HUSSON, H., FREEDMAN, A. S., CARDOSO, A. A., SCHULTZE, J., MUNOZ, O., STROLA, G., KUTOK, J., CARIDEO, E. G., DE BEAUMONT, R., CALIGARIS-CAPPIO, F. & GHIA, P. 2002. CXCL13 (BCA-1) is produced by follicular lymphoma cells: role in the accumulation of malignant B cells. *Br J Haematol,* 119**,** 492-5.

IRISH, J. M., CZERWINSKI, D. K., NOLAN, G. P. & LEVY, R. 2006. Altered B-cell receptor signaling kinetics distinguish human follicular lymphoma B cells from tumor-infiltrating nonmalignant B cells. *Blood,* 108**,** 3135-42.

IRISH, J. M., MYKLEBUST, J. H., ALIZADEH, A. A., HOUOT, R., SHARMAN, J. P., CZERWINSKI, D. K., NOLAN, G. P. & LEVY, R. 2010. B-cell signaling networks reveal a negative prognostic human lymphoma cell subset that emerges during tumor progression. *Proc Natl Acad Sci U S A,* 107**,** 12747-54.

JACOBS, H. & BROSS, L. 2001. Towards an understanding of somatic hypermutation. *Curr Opin Immunol,* 13**,** 208-18.

JAFFE, E. S., ZARATE-OSORNO, A., KINGMA, D. W., RAFFELD, M. & MEDEIROS, L. J. 1994. The interrelationship between Hodgkin's disease and non-Hodgkin's lymphomas. *Ann Oncol,* 5 Suppl 1**,** 7-11.

JAKUBZICK, C., GAUTIER, E. L., GIBBINGS, S. L., SOJKA, D. K., SCHLITZER, A., JOHNSON, T. E., IVANOV, S., DUAN, Q., BALA, S., CONDON, T., VAN ROOIJEN, N., GRAINGER, J. R., BELKAID, Y., MA'AYAN, A., RICHES, D. W., YOKOYAMA, W. M., GINHOUX, F., HENSON, P. M. & RANDOLPH, G. J. 2013. Minimal differentiation of classical monocytes as they survey steady-state tissues and transport antigen to lymph nodes. *Immunity,* 39**,** 599-610.

JEGALIAN, A. G., EBERLE, F. C., PACK, S. D., MIRVIS, M., RAFFELD, M., PITTALUGA, S. & JAFFE, E. S. 2011. Follicular lymphoma in situ: clinical implications and comparisons with partial involvement by follicular lymphoma. *Blood,* 118**,** 2976-84.

JIANG, Y., REDMOND, D., NIE, K., ENG, K. W., CLOZEL, T., MARTIN, P., TAN, L. H., MELNICK, A. M., TAM, W. & ELEMENTO, O. 2014. Deep sequencing reveals clonal evolution patterns and mutation events associated with relapse in B-cell lymphomas. *Genome Biol,* 15**,** 432.

JOHNSON, N. A., AL-TOURAH, A., BROWN, C. J., CONNORS, J. M., GASCOYNE, R. D. & HORSMAN, D. E. 2008. Prognostic significance of secondary cytogenetic alterations in follicular lymphomas. *Genes Chromosomes Cancer,* 47**,** 1038-48.

JUNLEN, H. R., PETERSON, S., KIMBY, E., LOCKMER, S., LINDEN, O., NILSSON-EHLE, H., ERLANSON, M., HAGBERG, H., RADLUND, A., HAGBERG, O. & WAHLIN, B. E. 2015. Follicular lymphoma in Sweden: nationwide improved survival in the rituximab era, particularly in elderly women: a Swedish Lymphoma Registry study. *Leukemia,* 29**,** 668-76.

KAHL, B. S. & YANG, D. T. 2016. Follicular lymphoma: evolving therapeutic strategies. *Blood,* 127**,** 2055-63.

KIRCHER, M., STENZEL, U. & KELSO, J. 2009. Improved base calling for the Illumina Genome Analyzer using machine learning strategies. *Genome Biol,* 10**,** R83.

KIRSCHBAUM, M., FRANKEL, P., POPPLEWELL, L., ZAIN, J., DELIOUKINA, M., PULLARKAT, V., MATSUOKA, D., PULONE, B., ROTTER, A. J., ESPINOZA-DELGADO, I., NADEMANEE, A., FORMAN, S. J., GANDARA, D. & NEWMAN, E. 2011. Phase II study of vorinostat for treatment of relapsed or refractory indolent non-Hodgkin's lymphoma and mantle cell lymphoma. *J Clin Oncol,* 29**,** 1198-203.

KISHIMOTO, W. & NISHIKORI, M. 2014. Molecular pathogenesis of follicular lymphoma. *J Clin Exp Hematop,* 54**,** 23-30.

KLEIN, J. L., NGUYEN, T. T., BIEN-WILLNER, G. A., CHEN, L., FOYIL, K. V., BARTLETT, N. L., DUNCAVAGE, E. J., HASSAN, A., FRATER, J. L. & KREISEL, F. 2014. CD163 immunohistochemistry is superior to CD68 in predicting outcome in classical Hodgkin lymphoma. *Am J Clin Pathol,* 141**,** 381-7.

KLEIN, U. & DALLA-FAVERA, R. 2008. Germinal centres: role in B-cell physiology and malignancy. *Nat Rev Immunol,* 8**,** 22-33.

KOSCHMIEDER, S., GOTTGENS, B., ZHANG, P., IWASAKI-ARAI, J., AKASHI, K., KUTOK, J. L., DAYARAM, T., GEARY, K., GREEN, A. R., TENEN, D. G. & HUETTNER, C. S. 2005. Inducible chronic phase of myeloid leukemia with expansion of hematopoietic stem cells in a transgenic model of BCR-ABL leukemogenesis. *Blood,* 105**,** 324-34.

KOSMIDIS, P., BONZHEIM, I., DUFKE, C., COLAK, S., HENTRICH, T., SCHROEDER, C., BAUER, P., ADAM, P. & FEND, F. 2017. Next generation sequencing of the clonal IGH rearrangement detects ongoing mutations and interfollicular trafficking in in situ follicular neoplasia. *PLoS One,* 12**,** e0178503.

KRIDEL, R., MOTTOK, A., FARINHA, P., BEN-NERIAH, S., ENNISHI, D., ZHENG, Y., CHAVEZ, E. A., SHULHA, H. P., TAN, K., CHAN, F. C., BOYLE, M., MEISSNER, B., TELENIUS, A., SEHN, L. H., MARRA, M. A., SHAH, S. P., STEIDL, C., CONNORS, J. M., SCOTT, D. W. & GASCOYNE, R. D. 2015a. Cell of origin of transformed follicular lymphoma. *Blood,* 126**,** 2118-27.

KRIDEL, R., XERRI, L., GELAS-DORE, B., TAN, K., FEUGIER, P., VAWDA, A., CANIONI, D., FARINHA, P., BOUSSETTA, S., MOCCIA, A. A., BRICE, P., CHAVEZ, E. A., KYLE, A. H., SCOTT, D. W., SANDERS, A. D., FABIANI, B., SLACK, G. W., MINCHINTON, A. I., HAIOUN, C., CONNORS, J. M., SEHN, L. H., STEIDL, C., GASCOYNE, R. D. & SALLES, G. 2015b. The Prognostic Impact of CD163-Positive Macrophages in Follicular Lymphoma: A Study from the BC Cancer Agency and the Lymphoma Study Association. *Clin Cancer Res,* 21**,** 3428-35.

KRIEG, C., HAN, P., STONE, R., GOULARTE, O. D. & KAYE, J. 2005. Functional analysis of B and T lymphocyte attenuator engagement on CD4+ and CD8+ T cells. *J Immunol,* 175**,** 6420-7.

KRYSIAK, K., GOMEZ, F., WHITE, B. S., MATLOCK, M., MILLER, C. A., TRANI, L., FRONICK, C. C., FULTON, R. S., KREISEL, F., CASHEN, A. F., CARSON, K. R., BERRIEN-ELLIOTT, M. M., BARTLETT, N. L., GRIFFITH, M., GRIFFITH, O. L. & FEHNIGER, T. A. 2017. Recurrent somatic mutations affecting B-cell receptor signaling pathway genes in follicular lymphoma. *Blood,* 129**,** 473-483.

KUPPERS, R. 2005. Mechanisms of B-cell lymphoma pathogenesis. *Nat Rev Cancer,* 5**,** 251-62.

KUPPERS, R., KLEIN, U., HANSMANN, M. L. & RAJEWSKY, K. 1999. Cellular origin of human B-cell lymphomas. *N Engl J Med,* 341**,** 1520-9.

KUPPERS, R., KLEIN, U., SCHWERING, I., DISTLER, V., BRAUNINGER, A., CATTORETTI, G., TU, Y., STOLOVITZKY, G. A., CALIFANO, A., HANSMANN, M. L. & DALLA-FAVERA, R. 2003. Identification of Hodgkin and Reed-Sternberg cell-specific genes by gene expression profiling. *J Clin Invest,* 111**,** 529-37.

KUPPERS, R., SOUSA, A. B., BAUR, A. S., STRICKLER, J. G., RAJEWSKY, K. & HANSMANN, M. L. 2001. Common germinal-center B-cell origin of the malignant cells in two composite lymphomas, involving classical Hodgkin's disease and either follicular lymphoma or B-CLL. *Mol Med,* 7**,** 285-92.

KUPPERS, R. & STEVENSON, F. K. 2018. Critical influences on the pathogenesis of follicular lymphoma. *Blood,* 131**,** 2297-2306.

LADETTO, M., DE MARCO, F., BENEDETTI, F., VITOLO, U., PATTI, C., RAMBALDI, A., PULSONI, A., MUSSO, M., LIBERATI, A. M., OLIVIERI, A., GALLAMINI, A., POGLIANI, E., ROTA SCALABRINI, D., CALLEA, V., DI RAIMONDO, F., PAVONE, V., TUCCI, A., CORTELAZZO, S., LEVIS, A., BOCCADORO, M., MAJOLINO, I., PILERI, A., GIANNI, A. M., PASSERA, R., CORRADINI, P. & TARELLA, C. 2008. Prospective, multicenter randomized GITMO/IIL trial comparing intensive (R-HDS) versus conventional (CHOP-R) chemoimmunotherapy in high-risk follicular lymphoma at diagnosis: the superior disease control of R-HDS does not translate into an overall survival advantage. *Blood,* 111**,** 4004-13.

LAUBACH, J. P., MOREAU, P., SAN-MIGUEL, J. F. & RICHARDSON, P. G. 2015. Panobinostat for the Treatment of Multiple Myeloma. *Clin Cancer Res,* 21**,** 4767-73.

LAUNAY, E., PANGAULT, C., BERTRAND, P., JARDIN, F., LAMY, T., TILLY, H., TARTE, K., BASTARD, C. & FEST, T. 2012. High rate of TNFRSF14 gene alterations related to 1p36 region in de novo follicular lymphoma and impact on prognosis. *Leukemia,* 26**,** 559-62.

LE GOUILL, S., DE GUIBERT, S., PLANCHE, L., BRICE, P., DUPUIS, J., CARTRON, G., VAN HOOF, A., CASASNOVAS, O., GYAN, E., TILLY, H., FRUCHART, C., DECONINCK, E., FITOUSSI, O., GASTAUD, L., DELWAIL, V., GABARRE, J., GRESSIN, R., BLANC, M., FOUSSARD, C. & SALLES, G. 2011. Impact of the use of autologous stem cell transplantation at first relapse both in naive and previously rituximab exposed follicular lymphoma patients treated in the GELA/GOELAMS FL2000 study. *Haematologica,* 96**,** 1128-35.

LEE, A. M., CLEAR, A. J., CALAMINICI, M., DAVIES, A. J., JORDAN, S., MACDOUGALL, F., MATTHEWS, J., NORTON, A. J., GRIBBEN, J. G., LISTER, T. A. & GOFF, L. K. 2006. Number of CD4+ cells and location of forkhead box protein P3-positive cells in diagnostic follicular lymphoma tissue microarrays correlates with outcome. *J Clin Oncol,* 24**,** 5052-9.

LEICH, E., ZAMO, A., HORN, H., HARALAMBIEVA, E., PUPPE, B., GASCOYNE, R. D., CHAN, W. C., BRAZIEL, R. M., RIMSZA, L. M., WEISENBURGER, D. D., DELABIE, J., JAFFE, E. S., FITZGIBBON, J., STAUDT, L. M., MUELLER-HERMELINK, H. K., CALAMINICI, M., CAMPO, E., OTT, G., HERNANDEZ, L. & ROSENWALD, A. 2011. MicroRNA profiles of t(14;18)-negative follicular lymphoma support a late germinal center B-cell phenotype. *Blood,* 118**,** 5550-8.

LENZ, G., DREYLING, M., SCHIEGNITZ, E., FORSTPOINTNER, R., WANDT, H., FREUND, M., HESS, G., TRUEMPER, L., DIEHL, V., KROPFF, M., KNEBA, M., SCHMITZ, N., METZNER, B., PFIRRMANN, M., UNTERHALT, M. & HIDDEMANN, W. 2004. Myeloablative radiochemotherapy followed by autologous stem cell transplantation in first remission prolongs progression-free survival in follicular lymphoma: results of a prospective, randomized trial of the German Low-Grade Lymphoma Study Group. *Blood,* 104**,** 2667-74.

LI, H., KAMINSKI, M. S., LI, Y., YILDIZ, M., OUILLETTE, P., JONES, S., FOX, H., JACOBI, K., SAIYA-CORK, K., BIXBY, D., LEBOVIC, D., ROULSTON, D., SHEDDEN, K., SABEL, M., MARENTETTE, L., CIMMINO, V., CHANG, A. E. & MALEK, S. N. 2014. Mutations in linker histone genes HIST1H1 B, C, D, and E; OCT2 (POU2F2); IRF8; and ARID1A underlying the pathogenesis of follicular lymphoma. *Blood,* 123**,** 1487-98.

LIMPENS, J., STAD, R., VOS, C., DE VLAAM, C., DE JONG, D., VAN OMMEN, G. J., SCHUURING, E. & KLUIN, P. M. 1995. Lymphoma-associated translocation t(14;18) in blood B cells of normal individuals. *Blood,* 85**,** 2528-36.

LINK, B. K., MAURER, M. J., NOWAKOWSKI, G. S., ANSELL, S. M., MACON, W. R., SYRBU, S. I., SLAGER, S. L., THOMPSON, C. A., INWARDS, D. J., JOHNSTON, P. B., COLGAN, J. P., WITZIG, T. E., HABERMANN, T. M. & CERHAN, J. R. 2013. Rates and outcomes of follicular lymphoma transformation in the immunochemotherapy era: a report from the University of Iowa/MayoClinic Specialized Program of Research Excellence Molecular Epidemiology Resource. *J Clin Oncol,* 31**,** 3272-8.

LINLEY, A., KRYSOV, S., PONZONI, M., JOHNSON, P. W., PACKHAM, G. & STEVENSON, F. K. 2015. Lectin binding to surface Ig variable regions provides a universal persistent activating signal for follicular lymphoma cells. *Blood,* 126**,** 1902-10.

LISO, A., CAPELLO, D., MARAFIOTI, T., TIACCI, E., CERRI, M., DISTLER, V., PAULLI, M., CARBONE, A., DELSOL, G., CAMPO, E., PILERI, S., PASQUALUCCI, L., GAIDANO, G.

& FALINI, B. 2006. Aberrant somatic hypermutation in tumor cells of nodular-lymphocyte-predominant and classic Hodgkin lymphoma. *Blood,* 108**,** 1013-20.

LIU, Q., SALAVERRIA, I., PITTALUGA, S., JEGALIAN, A. G., XI, L., SIEBERT, R., RAFFELD, M., HEWITT, S. M. & JAFFE, E. S. 2013. Follicular lymphomas in children and young adults: a comparison of the pediatric variant with usual follicular lymphoma. *Am J Surg Pathol,* 37**,** 333-43.

LIU, Y., HERNANDEZ, A. M., SHIBATA, D. & CORTOPASSI, G. A. 1994. BCL2 translocation frequency rises with age in humans. *Proc Natl Acad Sci U S A,* 91**,** 8910-4.

LO COCO, F., GAIDANO, G., LOUIE, D. C., OFFIT, K., CHAGANTI, R. S. & DALLA-FAVERA, R. 1993. p53 mutations are associated with histologic transformation of follicular lymphoma. *Blood,* 82**,** 2289-95.

LODER, F., MUTSCHLER, B., RAY, R. J., PAIGE, C. J., SIDERAS, P., TORRES, R., LAMERS, M. C. & CARSETTI, R. 1999. B cell development in the spleen takes place in discrete steps and is determined by the quality of B cell receptor-derived signals. *J Exp Med,* 190**,** 75-89.

LOPEZ-GIRAL, S., QUINTANA, N. E., CABRERIZO, M., ALFONSO-PEREZ, M., SALA-VALDES, M., DE SORIA, V. G., FERNANDEZ-RANADA, J. M., FERNANDEZ-RUIZ, E. & MUNOZ, C. 2004. Chemokine receptors that mediate B cell homing to secondary lymphoid tissues are highly expressed in B cell chronic lymphocytic leukemia and non-Hodgkin lymphomas with widespread nodular dissemination. *J Leukoc Biol,* 76**,** 462-71.

LORSBACH, R. B., SHAY-SEYMORE, D., MOORE, J., BANKS, P. M., HASSERJIAN, R. P., SANDLUND, J. T. & BEHM, F. G. 2002. Clinicopathologic analysis of follicular lymphoma occurring in children. *Blood,* 99**,** 1959-64.

LOSSOS, I. S., ALIZADEH, A. A., DIEHN, M., WARNKE, R., THORSTENSON, Y., OEFNER, P. J., BROWN, P. O., BOTSTEIN, D. & LEVY, R. 2002. Transformation of follicular lymphoma to diffuse large-cell lymphoma: alternative patterns with increased or decreased expression of c-myc and its regulated genes. *Proc Natl Acad Sci U S A,* 99**,** 8886-91.

LOSSOS, I. S. & GASCOYNE, R. D. 2011. Transformation of follicular lymphoma. *Best Pract Res Clin Haematol,* 24**,** 147-63.

LOUISSAINT, A., JR., ACKERMAN, A. M., DIAS-SANTAGATA, D., FERRY, J. A., HOCHBERG, E. P., HUANG, M. S., IAFRATE, A. J., LARA, D. O., PINKUS, G. S., SALAVERRIA, I., SIDDIQUEE, Z., SIEBERT, R., WEINSTEIN, H. J., ZUKERBERG, L. R., HARRIS, N. L. & HASSERJIAN, R. P. 2012. Pediatric-type nodal follicular lymphoma: an indolent clonal proliferation in children and adults with high proliferation index and no BCL2 rearrangement. *Blood,* 120**,** 2395-404.

LOUISSAINT, A., JR., SCHAFERNAK, K. T., GEYER, J. T., KOVACH, A. E., GHANDI, M., GRATZINGER, D., ROTH, C. G., PAXTON, C. N., KIM, S., NAMGYAL, C., MORIN, R., MORGAN, E. A., NEUBERG, D. S., SOUTH, S. T., HARRIS, M. H., HASSERJIAN, R. P., HOCHBERG, E. P., GARRAWAY, L. A., HARRIS, N. L. & WEINSTOCK, D. M. 2016. Pediatric-type nodal follicular lymphoma: a biologically distinct lymphoma with frequent MAPK pathway mutations. *Blood,* 128**,** 1093-100.

LOWRY, L., SMITH, P., QIAN, W., FALK, S., BENSTEAD, K., ILLIDGE, T., LINCH, D., ROBINSON, M., JACK, A. & HOSKIN, P. 2011. Reduced dose radiotherapy for local control in non-Hodgkin lymphoma: a randomised phase III trial. *Radiother Oncol,* 100**,** 86-92.

LU, X., NECHUSHTAN, H., DING, F., ROSADO, M. F., SINGAL, R., ALIZADEH, A. A. & LOSSOS, I. S. 2005. Distinct IL-4-induced gene expression, proliferation, and intracellular signaling in germinal center B-cell-like and activated B-cell-like diffuse large-cell lymphomas. *Blood,* 105**,** 2924-32.

LYDYARD, P. M., JEWELL, A. P., JAMIN, C. & YOUINOU, P. Y. 1999. CD5 B cells and B-cell malignancies. *Curr Opin Hematol,* 6**,** 30-6.

MAMESSIER, E., DREVET, C., BROUSSAIS-GUILLAUMOT, F., MOLLICHELLA, M. L., GARCIAZ, S., ROULLAND, S., BENCHETRIT, M., NADEL, B. & XERRI, L. 2015. Contiguous follicular lymphoma and follicular lymphoma in situ harboring N-glycosylated sites. *Haematologica,* 100**,** e155-7.

MAMESSIER, E., SONG, J. Y., EBERLE, F. C., PACK, S., DREVET, C., CHETAILLE, B., ABDULLAEV, Z., ADELAIDE, J., BIRNBAUM, D., CHAFFANET, M., PITTALUGA, S., ROULLAND, S., CHOTT, A., JAFFE, E. S. & NADEL, B. 2014. Early lesions of follicular lymphoma: a genetic perspective. *Haematologica,* 99**,** 481-8.

MARAFIOTI, T., HUMMEL, M., ANAGNOSTOPOULOS, I., FOSS, H. D., HUHN, D. & STEIN, H. 1999. Classical Hodgkin's disease and follicular lymphoma originating from the same germinal center B cell. *J Clin Oncol,* 17**,** 3804-9.

MARCUS, R., IMRIE, K., BELCH, A., CUNNINGHAM, D., FLORES, E., CATALANO, J., SOLAL-CELIGNY, P., OFFNER, F., WALEWSKI, J., RAPOSO, J., JACK, A. & SMITH, P. 2005. CVP chemotherapy plus rituximab compared with CVP as first-line treatment for advanced follicular lymphoma. *Blood,* 105**,** 1417-23.

MARCUS, R., IMRIE, K., SOLAL-CELIGNY, P., CATALANO, J. V., DMOSZYNSKA, A., RAPOSO, J. C., OFFNER, F. C., GOMEZ-CODINA, J., BELCH, A., CUNNINGHAM, D., WASSNER-FRITSCH, E. & STEIN, G. 2008. Phase III study of R-CVP compared with cyclophosphamide, vincristine, and prednisone alone in patients with previously untreated advanced follicular lymphoma. *J Clin Oncol,* 26**,** 4579-86.

MARTIN-GUERRERO, I., SALAVERRIA, I., BURKHARDT, B., SZCZEPANOWSKI, M., BAUDIS, M., BENS, S., DE LEVAL, L., GARCIA-ORAD, A., HORN, H., LISFELD, J., PELLISSERY, S., KLAPPER, W., OSCHLIES, I. & SIEBERT, R. 2013. Recurrent loss of heterozygosity in 1p36 associated with TNFRSF14 mutations in IRF4 translocation negative pediatric follicular lymphomas. *Haematologica,* 98**,** 1237-41.

MARTIN, F. & KEARNEY, J. F. 2000. Positive selection from newly formed to marginal zone B cells depends on the rate of clonal production, CD19, and btk. *Immunity,* 12**,** 39-49.

MARTINEZ-CLIMENT, J. A., ALIZADEH, A. A., SEGRAVES, R., BLESA, D., RUBIO-MOSCARDO, F., ALBERTSON, D. G., GARCIA-CONDE, J., DYER, M. J., LEVY, R., PINKEL, D. & LOSSOS, I. S. 2003. Transformation of follicular lymphoma to diffuse large cell lymphoma is associated with a heterogeneous set of DNA copy number and gene expression alterations. *Blood,* 101**,** 3109-17.

MCCANN, K. J., JOHNSON, P. W., STEVENSON, F. K. & OTTENSMEIER, C. H. 2006. Universal N-glycosylation sites introduced into the B-cell receptor of follicular lymphoma by somatic mutation: a second tumorigenic event? *Leukemia,* 20**,** 530-4.

MCCANN, K. J., OTTENSMEIER, C. H., CALLARD, A., RADCLIFFE, C. M., HARVEY, D. J., DWEK, R. A., RUDD, P. M., SUTTON, B. J., HOBBY, P. & STEVENSON, F. K. 2008. Remarkable selective glycosylation of the immunoglobulin variable region in follicular lymphoma. *Mol Immunol,* 45**,** 1567-72.

MCDONNELL, T. J. & KORSMEYER, S. J. 1991. Progression from lymphoid hyperplasia to high-grade malignant lymphoma in mice transgenic for the t(14; 18). *Nature,* 349**,** 254-6.

MCGRANAHAN, N. & SWANTON, C. 2017. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. *Cell,* 168**,** 613-628.

MCGREAL, E. P., MARTINEZ-POMARES, L. & GORDON, S. 2004. Divergent roles for C-type lectins expressed by cells of the innate immune system. *Mol Immunol,* 41**,** 1109-21.

MERLO, A., HERMAN, J. G., MAO, L., LEE, D. J., GABRIELSON, E., BURGER, P. C., BAYLIN, S. B. & SIDRANSKY, D. 1995. 5' CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers. *Nat Med,* 1**,** 686-92.

MOLAVI, O., WANG, P., ZAK, Z., GELEBART, P., BELCH, A. & LAI, R. 2013. Gene methylation and silencing of SOCS3 in mantle cell lymphoma. *Br J Haematol,* 161**,** 348-56.

MONTOTO, S., CANALS, C., ROHATINER, A. Z., TAGHIPOUR, G., SUREDA, A., SCHMITZ, N., GISSELBRECHT, C., FOUILLARD, L., MILPIED, N., HAIOUN, C., SLAVIN, S., CONDE, E., FRUCHART, C., FERRANT, A., LEBLOND, V., TILLY, H., LISTER, T. A. & GOLDSTONE, A. H. 2007a. Long-term follow-up of high-dose treatment with autologous haematopoietic progenitor cell support in 693 patients with follicular lymphoma: an EBMT registry study. *Leukemia,* 21**,** 2324-31.

MONTOTO, S., DAVIES, A. J., MATTHEWS, J., CALAMINICI, M., NORTON, A. J., AMESS, J., VINNICOMBE, S., WATERS, R., ROHATINER, A. Z. & LISTER, T. A. 2007b. Risk and clinical implications of transformation of follicular lymphoma to diffuse large B-cell lymphoma. *J Clin Oncol,* 25**,** 2426-33.

MORIN, R. D., JOHNSON, N. A., SEVERSON, T. M., MUNGALL, A. J., AN, J., GOYA, R., PAUL, J. E., BOYLE, M., WOOLCOCK, B. W., KUCHENBAUER, F., YAP, D., HUMPHRIES, R. K., GRIFFITH, O. L., SHAH, S., ZHU, H., KIMBARA, M., SHASHKIN, P., CHARLOT, J. F., TCHERPAKOV, M., CORBETT, R., TAM, A., VARHOL, R., SMAILUS, D., MOKSA, M., ZHAO, Y., DELANEY, A., QIAN, H., BIROL, I., SCHEIN, J., MOORE, R., HOLT, R., HORSMAN, D. E., CONNORS, J. M., JONES, S., APARICIO, S., HIRST, M., GASCOYNE, R. D. & MARRA, M. A. 2010. Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nat Genet,* 42**,** 181-5.

MORIN, R. D., MENDEZ-LAGO, M., MUNGALL, A. J., GOYA, R., MUNGALL, K. L., CORBETT, R. D., JOHNSON, N. A., SEVERSON, T. M., CHIU, R., FIELD, M., JACKMAN, S., KRZYWINSKI, M., SCOTT, D. W., TRINH, D. L., TAMURA-WELLS, J., LI, S., FIRME, M. R., ROGIC, S., GRIFFITH, M., CHAN, S., YAKOVENKO, O., MEYER, I. M., ZHAO, E. Y., SMAILUS, D., MOKSA, M., CHITTARANJAN, S., RIMSZA, L., BROOKS-WILSON, A., SPINELLI, J. J., BEN-NERIAH, S., MEISSNER, B., WOOLCOCK, B., BOYLE, M., MCDONALD, H., TAM, A., ZHAO, Y., DELANEY, A., ZENG, T., TSE, K., BUTTERFIELD, Y., BIROL, I., HOLT, R., SCHEIN, J., HORSMAN, D. E., MOORE, R., JONES, S. J., CONNORS, J. M., HIRST, M., GASCOYNE, R. D. & MARRA, M. A. 2011. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature,* 476**,** 298-303.

MORSCHHAUSER, F., SALLES, G., MCKAY, P., TILLY, H., SCHMITT, A., GERECITANO, J., JOHNSON, P., LE GOUILL, S., DICKINSON, M. J., FRUCHART, C., LAMY, T., CHAIDOS, A., JURCZAK, W., OPAT, S., RADFORD, J., ZINZANI, P. L., ASSOULINE, S., CARTRON, G., CLAWSON, A., PICAZIO, N., RIBICH, S., BLAKEMORE, S. J., LARUS, J., MIAO, H., HO, P. T. & RIBRAG, V. 2017. INTERIM REPORT FROM A PHASE 2 MULTICENTER STUDY OF TAZEMETOSTAT, AN EZH2 INHIBITOR, IN PATIENTS WITH RELAPSED OR REFRACTORY B-CELL NON-HODGKIN LYMPHOMAS. *Hematological Oncology,* 35**,** 24-25.

MORSCHHAUSER, F., SEYMOUR, J., FEUGIER, P., OFFNER, F., LOPEZ-GUILLERMO, A., BOUABDALLAH, R., PEDERSEN, L. & ET, A. 2011. Impact of Induction Chemotherapy Regimen on Response, Safety and Outcome in the PRIMA Study. *Annals of Oncology,* 22**,** 89.

MOUNIER, M., BOSSARD, N., REMONTET, L., BELOT, A., MINICOZZI, P., DE ANGELIS, R., CAPOCACCIA, R., IWAZ, J., MONNEREAU, A., TROUSSARD, X., SANT, M., MAYNADIE, M. & GIORGI, R. 2015. Changes in dynamics of excess mortality rates

and net survival after diagnosis of follicular lymphoma or diffuse large B-cell lymphoma: comparison between European population-based data (EUROCARE-5). *Lancet Haematol,* 2**,** e481-91.

NANJANGUD, G., RAO, P. H., TERUYA-FELDSTEIN, J., DONNELLY, G., QIN, J., MEHRA, S., JHANWAR, S. C., ZELENETZ, A. D. & CHAGANTI, R. S. 2007. Molecular cytogenetic analysis of follicular lymphoma (FL) provides detailed characterization of chromosomal instability associated with the t(14;18)(q32;q21) positive and negative subsets and histologic progression. *Cytogenet Genome Res,* 118**,** 337-44.

NATKUNAM, Y., SOSLOW, R., MATOLCSY, A., DOLEZAL, M., BHARGAVA, V., KNOWLES, D. M. & WARNKE, R. 2004. Immunophenotypic and genotypic characterization of progression in follicular lymphomas. *Appl Immunohistochem Mol Morphol,* 12**,** 97-104.

NIIRO, H. & CLARK, E. A. 2002. Regulation of B-cell fate by antigen-receptor signals. *Nat Rev Immunol,* 2**,** 945-56.

NISHIZUMI, H., HORIKAWA, K., MLINARIC-RASCAN, I. & YAMAMOTO, T. 1998. A double-edged kinase Lyn: a positive and negative regulator for antigen receptor-mediated signals. *J Exp Med,* 187**,** 1343-8.

NONAKA, M., MA, B. Y., MURAI, R., NAKAMURA, N., BABA, M., KAWASAKI, N., HODOHARA, K., ASANO, S. & KAWASAKI, T. 2008. Glycosylation-dependent interactions of C-type lectin DC-SIGN with colorectal tumor-associated Lewis glycans impair the function and differentiation of monocyte-derived dendritic cells. *J Immunol,* 180**,** 3347-56.

OFFIT, K., PARSA, N. Z., GAIDANO, G., FILIPPA, D. A., LOUIE, D., PAN, D., JHANWAR, S. C., DALLA-FAVERA, R. & CHAGANTI, R. S. 1993. 6q deletions define distinct clinico-pathologic subsets of non-Hodgkin's lymphoma. *Blood,* 82**,** 2157-62.

OGURA, M., ANDO, K., SUZUKI, T., ISHIZAWA, K., OH, S. Y., ITOH, K., YAMAMOTO, K., AU, W. Y., TIEN, H. F., MATSUNO, Y., TERAUCHI, T., YAMAMOTO, K., MORI, M., TANAKA, Y., SHIMAMOTO, T., TOBINAI, K. & KIM, W. S. 2014. A multicentre phase II study of vorinostat in patients with relapsed or refractory indolent B-cell non-Hodgkin lymphoma and mantle cell lymphoma. *Br J Haematol,* 165**,** 768-76.

OKOSUN, J., BODOR, C., WANG, J., ARAF, S., YANG, C. Y., PAN, C., BOLLER, S., CITTARO, D., BOZEK, M., IQBAL, S., MATTHEWS, J., WRENCH, D., MARZEC, J., TAWANA, K., POPOV, N., O'RIAIN, C., O'SHEA, D., CARLOTTI, E., DAVIES, A., LAWRIE, C. H., MATOLCSY, A., CALAMINICI, M., NORTON, A., BYERS, R. J., MEIN, C., STUPKA, E., LISTER, T. A., LENZ, G., MONTOTO, S., GRIBBEN, J. G., FAN, Y., GROSSCHEDL, R., CHELALA, C. & FITZGIBBON, J. 2014. Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. *Nat Genet,* 46**,** 176-181.

ORICCHIO, E., NANJANGUD, G., WOLFE, A. L., SCHATZ, J. H., MAVRAKIS, K. J., JIANG, M., LIU, X., BRUNO, J., HEGUY, A., OLSHEN, A. B., SOCCI, N. D., TERUYA-FELDSTEIN, J., WEIS-GARCIA, F., TAM, W., SHAKNOVICH, R., MELNICK, A., HIMANEN, J. P., CHAGANTI, R. S. & WENDEL, H. G. 2011. The Eph-receptor A7 is a soluble tumor suppressor for follicular lymphoma. *Cell,* 147**,** 554-64.

OSCIER, D., WADE, R., DAVIS, Z., MORILLA, A., BEST, G., RICHARDS, S., ELSE, M., MATUTES, E. & CATOVSKY, D. 2010. Prognostic factors identified three risk groups in the LRF CLL4 trial, independent of treatment allocation. *Haematologica,* 95**,** 1705-12.

OTSUKI, T., YANO, T., CLARK, H. M., BASTARD, C., KERCKAERT, J. P., JAFFE, E. S. & RAFFELD, M. 1995. Analysis of LAZ3 (BCL-6) status in B-cell non-Hodgkin's lymphomas: results of rearrangement and gene expression studies and a mutational analysis of coding region sequences. *Blood,* 85**,** 2877-84.

PANGAULT, C., AME-THOMAS, P., RUMINY, P., ROSSILLE, D., CARON, G., BAIA, M., DE VOS, J., ROUSSEL, M., MONVOISIN, C., LAMY, T., TILLY, H., GAULARD, P., TARTE, K. & FEST, T. 2010. Follicular lymphoma cell niche: identification of a preeminent IL-4-dependent T(FH)-B cell axis. *Leukemia,* 24**,** 2080-9.

PANI, G., KOZLOWSKI, M., CAMBIER, J. C., MILLS, G. B. & SIMINOVITCH, K. A. 1995. Identification of the tyrosine phosphatase PTP1C as a B cell antigen receptor-associated protein involved in the regulation of B cell signaling. *J Exp Med,* 181**,** 2077-84.

PASPARAKIS, M., ALEXOPOULOU, L., DOUNI, E. & KOLLIAS, G. 1996. Tumour necrosis factors in immune regulation: everything that's interesting is...new! *Cytokine Growth Factor Rev,* 7**,** 223-9.

PASQUALUCCI, L., DOMINGUEZ-SOLA, D., CHIARENZA, A., FABBRI, G., GRUNN, A., TRIFONOV, V., KASPER, L. H., LERACH, S., TANG, H., MA, J., ROSSI, D., CHADBURN, A., MURTY, V. V., MULLIGHAN, C. G., GAIDANO, G., RABADAN, R., BRINDLE, P. K. & DALLA-FAVERA, R. 2011a. Inactivating mutations of acetyltransferase genes in B-cell lymphoma. *Nature,* 471**,** 189-95.

PASQUALUCCI, L., KHIABANIAN, H., FANGAZIO, M., VASISHTHA, M., MESSINA, M., HOLMES, A. B., OUILLETTE, P., TRIFONOV, V., ROSSI, D., TABBO, F., PONZONI, M., CHADBURN, A., MURTY, V. V., BHAGAT, G., GAIDANO, G., INGHIRAMI, G., MALEK, S. N., RABADAN, R. & DALLA-FAVERA, R. 2014. Genetics of follicular lymphoma transformation. *Cell Rep,* 6**,** 130-40.

PASQUALUCCI, L., NEUMEISTER, P., GOOSSENS, T., NANJANGUD, G., CHAGANTI, R. S., KUPPERS, R. & DALLA-FAVERA, R. 2001. Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature,* 412**,** 341-6.

PASQUALUCCI, L., TRIFONOV, V., FABBRI, G., MA, J., ROSSI, D., CHIARENZA, A., WELLS, V. A., GRUNN, A., MESSINA, M., ELLIOT, O., CHAN, J., BHAGAT, G., CHADBURN, A., GAIDANO, G., MULLIGHAN, C. G., RABADAN, R. & DALLA-FAVERA, R. 2011b. Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat Genet,* 43**,** 830-7.

PINTO, A., HUTCHISON, R. E., GRANT, L. H., TREVENEN, C. L. & BERARD, C. W. 1990. Follicular lymphomas in pediatric patients. *Mod Pathol,* 3**,** 308-13.

RADCLIFFE, C. M., ARNOLD, J. N., SUTER, D. M., WORMALD, M. R., HARVEY, D. J., ROYLE, L., MIMURA, Y., KIMURA, Y., SIM, R. B., INOGES, S., RODRIGUEZ-CALVILLO, M., ZABALEGUI, N., DE CERIO, A. L., POTTER, K. N., MOCKRIDGE, C. I., DWEK, R. A., BENDANDI, M., RUDD, P. M. & STEVENSON, F. K. 2007. Human follicular lymphoma cells contain oligomannose glycans in the antigen-binding site of the B-cell receptor. *J Biol Chem,* 282**,** 7405-15.

RAMSAY, A. G., CLEAR, A. J., KELLY, G., FATAH, R., MATTHEWS, J., MACDOUGALL, F., LISTER, T. A., LEE, A. M., CALAMINICI, M. & GRIBBEN, J. G. 2009. Follicular lymphoma cells induce T-cell immunologic synapse dysfunction that can be repaired with lenalidomide: implications for the tumor microenvironment and immunotherapy. *Blood,* 114**,** 4713-20.

RAMSAY, A. G. & GRIBBEN, J. G. 2011. The kiss of death in FL. *Blood,* 118**,** 5365-6.

RAWAL, S., CHU, F., ZHANG, M., PARK, H. J., NATTAMAI, D., KANNAN, S., SHARMA, R., DELGADO, D., CHOU, T., LIN, H. Y., BALADANDAYUTHAPANI, V., LUONG, A., VEGA, F., FOWLER, N., DONG, C., DAVIS, R. E. & NEELAPU, S. S. 2013. Cross talk between follicular Th cells and tumor cells in human follicular lymphoma promotes immune evasion in the tumor microenvironment. *J Immunol,* 190**,** 6681-93.

REIK, W. 2007. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature,* 447**,** 425-32.

RISDALL, R., HOPPE, R. T. & WARNKE, R. 1979. Non-Hodgkin's lymphoma: a study of the evolution of the disease based upon 92 autopsied cases. *Cancer,* 44**,** 529-42.

ROBINSON, K. S., WILLIAMS, M. E., VAN DER JAGT, R. H., COHEN, P., HERST, J. A., TULPULE, A., SCHWARTZBERG, L. S., LEMIEUX, B. & CHESON, B. D. 2008. Phase II multicenter study of bendamustine plus rituximab in patients with relapsed indolent B-cell and mantle cell non-Hodgkin's lymphoma. *J Clin Oncol,* 26**,** 4473-9.

ROBINSON, M. J., SANCHO, D., SLACK, E. C., LEIBUNDGUT-LANDMANN, S. & REIS E SOUSA, C. 2006. Myeloid C-type lectins in innate immunity. *Nat Immunol,* 7**,** 1258-65.

ROLLI, V., GALLWITZ, M., WOSSNING, T., FLEMMING, A., SCHAMEL, W. W., ZURN, C. & RETH, M. 2002. Amplification of B cell antigen receptor signaling by a Syk/ITAM positive feedback loop. *Mol Cell,* 10**,** 1057-69.

ROOS, J., DIGREGORIO, P. J., YEROMIN, A. V., OHLSEN, K., LIOUDYNO, M., ZHANG, S., SAFRINA, O., KOZAK, J. A., WAGNER, S. L., CAHALAN, M. D., VELICELEBI, G. & STAUDERMAN, K. A. 2005. STIM1, an essential and conserved component of store-operated Ca2+ channel function. *J Cell Biol,* 169**,** 435-45.

ROSSI, D., CERRI, M., CAPELLO, D., DEAMBROGI, C., BERRA, E., FRANCESCHETTI, S., ALABISO, O., GLOGHINI, A., PAULLI, M., CARBONE, A., PILERI, S. A., PASQUALUCCI, L. & GAIDANO, G. 2005. Aberrant somatic hypermutation in primary mediastinal large B-cell lymphoma. *Leukemia,* 19**,** 2363-6.

ROULLAND, S., KELLY, R. S., MORGADO, E., SUNGALEE, S., SOLAL-CELIGNY, P., COLOMBAT, P., JOUVE, N., PALLI, D., PALA, V., TUMINO, R., PANICO, S., SACERDOTE, C., QUIROS, J. R., GONZALES, C. A., SANCHEZ, M. J., DORRONSORO, M., NAVARRO, C., BARRICARTE, A., TJONNELAND, A., OLSEN, A., OVERVAD, K., CANZIAN, F., KAAKS, R., BOEING, H., DROGAN, D., NIETERS, A., CLAVEL-CHAPELON, F., TRICHOPOULOU, A., TRICHOPOULOS, D., LAGIOU, P., BUENO-DE-MESQUITA, H. B., PEETERS, P. H., VERMEULEN, R., HALLMANS, G., MELIN, B., BORGQUIST, S., CARLSON, J., LUND, E., WEIDERPASS, E., KHAW, K. T., WAREHAM, N., KEY, T. J., TRAVIS, R. C., FERRARI, P., ROMIEU, I., RIBOLI, E., SALLES, G., VINEIS, P. & NADEL, B. 2014. t(14;18) Translocation: A predictive blood biomarker for follicular lymphoma. *J Clin Oncol,* 32**,** 1347-55.

ROULLAND, S., LEBAILLY, P., LECLUSE, Y., BRIAND, M., POTTIER, D. & GAUDUCHON, P. 2004. Characterization of the t(14;18) BCL2-IGH translocation in farmers occupationally exposed to pesticides. *Cancer Res,* 64**,** 2264-9.

ROULLAND, S., LEBAILLY, P., LECLUSE, Y., HEUTTE, N., NADEL, B. & GAUDUCHON, P. 2006a. Long-term clonal persistence and evolution of t(14;18)-bearing B cells in healthy individuals. *Leukemia,* 20**,** 158-62.

ROULLAND, S., NAVARRO, J. M., GRENOT, P., MILILI, M., AGOPIAN, J., MONTPELLIER, B., GAUDUCHON, P., LEBAILLY, P., SCHIFF, C. & NADEL, B. 2006b. Follicular lymphoma-like B cells in healthy individuals: a novel intermediate step in early lymphomagenesis. *J Exp Med,* 203**,** 2425-31.

RULAND, J. & MAK, T. W. 2003. Transducing signals from antigen receptors to nuclear factor kappaB. *Immunol Rev,* 193**,** 93-100.

RUMMEL, M. J., AL-BATRAN, S. E., KIM, S. Z., WELSLAU, M., HECKER, R., KOFAHL-KRAUSE, D., JOSTEN, K. M., DURK, H., ROST, A., NEISE, M., VON GRUNHAGEN, U., CHOW, K. U., HANSMANN, M. L., HOELZER, D. & MITROU, P. S. 2005. Bendamustine plus rituximab is effective and has a favorable toxicity profile in the treatment of mantle cell and low-grade non-Hodgkin's lymphoma. *J Clin Oncol,* 23**,** 3383-9.

RUMMEL, M. J., NIEDERLE, N., MASCHMEYER, G., BANAT, G. A., VON GRUNHAGEN, U., LOSEM, C., KOFAHL-KRAUSE, D., HEIL, G., WELSLAU, M., BALSER, C., KAISER, U., WEIDMANN, E., DURK, H., BALLO, H., STAUCH, M., ROLLER, F., BARTH, J., HOELZER, D., HINKE, A. & BRUGGER, W. 2013. Bendamustine plus rituximab

versus CHOP plus rituximab as first-line treatment for patients with indolent and mantle-cell lymphomas: an open-label, multicentre, randomised, phase 3 non-inferiority trial. *Lancet,* 381**,** 1203-10.

SABOURI, Z., SCHOFIELD, P., HORIKAWA, K., SPIERINGS, E., KIPLING, D., RANDALL, K. L., LANGLEY, D., ROOME, B., VAZQUEZ-LOMBARDI, R., ROUET, R., HERMES, J., CHAN, T. D., BRINK, R., DUNN-WALTERS, D. K., CHRIST, D. & GOODNOW, C. C. 2014. Redemption of autoantibodies on anergic B cells by variable-region glycosylation and mutation away from self-reactivity. *Proc Natl Acad Sci U S A,* 111**,** E2567-75.

SACHEN, K. L., STROHMAN, M. J., SINGLETARY, J., ALIZADEH, A. A., KATTAH, N. H., LOSSOS, C., MELLINS, E. D., LEVY, S. & LEVY, R. 2012. Self-antigen recognition by follicular lymphoma B-cell receptors. *Blood,* 120**,** 4182-90.

SALLES, G., SEYMOUR, J. F., OFFNER, F., LOPEZ-GUILLERMO, A., BELADA, D., XERRI, L., FEUGIER, P., BOUABDALLAH, R., CATALANO, J. V., BRICE, P., CABALLERO, D., HAIOUN, C., PEDERSEN, L. M., DELMER, A., SIMPSON, D., LEPPA, S., SOUBEYRAN, P., HAGENBEEK, A., CASASNOVAS, O., INTRAGUMTORNCHAI, T., FERME, C., DA SILVA, M. G., SEBBAN, C., LISTER, A., ESTELL, J. A., MILONE, G., SONET, A., MENDILA, M., COIFFIER, B. & TILLY, H. 2011. Rituximab maintenance for 2 years in patients with high tumour burden follicular lymphoma responding to rituximab plus chemotherapy (PRIMA): a phase 3, randomised controlled trial. *Lancet,* 377**,** 42-51.

SANDER, C. A., YANO, T., CLARK, H. M., HARRIS, C., LONGO, D. L., JAFFE, E. S. & RAFFELD, M. 1993. p53 mutation is associated with progression in follicular lymphomas. *Blood,* 82**,** 1994-2004.

SAXON, E. & BERTOZZI, C. R. 2001. Chemical and biological strategies for engineering cell surface glycosylation. *Annu Rev Cell Dev Biol,* 17**,** 1-23.

SCHERER, F., NAVARRETE, M. A., BERTINETTI-LAPATKI, C., BOEHM, J., SCHMITT-GRAEFF, A. & VEELKEN, H. 2016. Isotype-switched follicular lymphoma displays dissociation between activation-induced cytidine deaminase expression and somatic hypermutation. *Leuk Lymphoma,* 57**,** 151-60.

SCHMATZ, A. I., STREUBEL, B., KRETSCHMER-CHOTT, E., PUSPOK, A., JAGER, U., MANNHALTER, C., TIEMANN, M., OTT, G., FISCHBACH, W., HERZOG, P., SEITZ, G., STOLTE, M., RADERER, M. & CHOTT, A. 2011. Primary follicular lymphoma of the duodenum is a distinct mucosal/submucosal variant of follicular lymphoma: a retrospective study of 63 cases. *J Clin Oncol,* 29**,** 1445-51.

SCHMIDT, J., GONG, S., MARAFIOTI, T., MANKEL, B., GONZALEZ-FARRE, B., BALAGUE, O., MOZOS, A., CABECADAS, J., VAN DER WALT, J., HOEHN, D., ROSENWALD, A., OTT, G., DOJCINOV, S., EGAN, C., NADEU, F., RAMIS-ZALDIVAR, J. E., CLOT, G., BARCENA, C., PEREZ-ALONSO, V., ENDRIS, V., PENZEL, R., LOME-MALDONADO, C., BONZHEIM, I., FEND, F., CAMPO, E., JAFFE, E. S., SALAVERRIA, I. & QUINTANILLA-MARTINEZ, L. 2016. Genome-wide analysis of pediatric-type follicular lymphoma reveals low genetic complexity and recurrent alterations of TNFRSF14 gene. *Blood,* 128**,** 1101-11.

SCHMIDT, J., SALAVERRIA, I., HAAKE, A., BONZHEIM, I., ADAM, P., MONTES-MORENO, S., PIRIS, M. A., FEND, F., SIEBERT, R. & QUINTANILLA-MARTINEZ, L. 2014. Increasing genomic and epigenomic complexity in the clonal evolution from in situ to manifest t(14;18)-positive follicular lymphoma. *Leukemia,* 28**,** 1103-12.

SCHMITZ, R., CERIBELLI, M., PITTALUGA, S., WRIGHT, G. & STAUDT, L. M. 2014. Oncogenic mechanisms in Burkitt lymphoma. *Cold Spring Harb Perspect Med,* 4.

SCHNEIDER, D., DUHREN-VON MINDEN, M., ALKHATIB, A., SETZ, C., VAN BERGEN, C. A., BENKISSER-PETERSEN, M., WILHELM, I., VILLRINGER, S., KRYSOV, S., PACKHAM, G., ZIRLIK, K., ROMER, W., BUSKE, C., STEVENSON, F. K., VEELKEN, H. & JUMAA, H.

2015. Lectins from opportunistic bacteria interact with acquired variable-region glycans of surface immunoglobulin in follicular lymphoma. *Blood,* 125**,** 3287-96.

SCHOUTEN, H. C., QIAN, W., KVALOY, S., PORCELLINI, A., HAGBERG, H., JOHNSEN, H. E., DOORDUIJN, J. K., SYDES, M. R. & KVALHEIM, G. 2003. High-dose therapy improves progression-free survival and survival in relapsed follicular non-Hodgkin's lymphoma: results from the randomized European CUP trial. *J Clin Oncol,* 21**,** 3918-27.

SCHROEDER, H. W., JR. & DIGHIERO, G. 1994. The pathogenesis of chronic lymphocytic leukemia: analysis of the antibody repertoire. *Immunol Today,* 15**,** 288-94.

SCHULER, F., DOLKEN, L., HIRT, C., KIEFER, T., BERG, T., FUSCH, G., WEITMANN, K., HOFFMANN, W., FUSCH, C., JANZ, S., RABKIN, C. S. & DOLKEN, G. 2009. Prevalence and frequency of circulating t(14;18)-MBR translocation carrying cells in healthy individuals. *Int J Cancer,* 124**,** 958-63.

SCHWAENEN, C., VIARDOT, A., BERGER, H., BARTH, T. F., BENTINK, S., DOHNER, H., ENZ, M., FELLER, A. C., HANSMANN, M. L., HUMMEL, M., KESTLER, H. A., KLAPPER, W., KREUZ, M., LENZE, D., LOEFFLER, M., MOLLER, P., MULLER-HERMELINK, H. K., OTT, G., ROSOLOWSKI, M., ROSENWALD, A., RUF, S., SIEBERT, R., SPANG, R., STEIN, H., TRUEMPER, L., LICHTER, P., BENTZ, M. & WESSENDORF, S. 2009. Microarray-based genomic profiling reveals novel genomic aberrations in follicular lymphoma which associate with patient survival and gene expression status. *Genes Chromosomes Cancer,* 48**,** 39-54.

SEBBAN, C., BRICE, P., DELARUE, R., HAIOUN, C., SOULEAU, B., MOUNIER, N., BROUSSE, N., FEUGIER, P., TILLY, H., SOLAL-CELIGNY, P. & COIFFIER, B. 2008. Impact of rituximab and/or high-dose therapy with autotransplant at time of relapse in patients with follicular lymphoma: a GELA study. *J Clin Oncol,* 26**,** 3614-20.

SEDY, J. R., GAVRIELI, M., POTTER, K. G., HURCHLA, M. A., LINDSLEY, R. C., HILDNER, K., SCHEU, S., PFEFFER, K., WARE, C. F., MURPHY, T. L. & MURPHY, K. M. 2005. B and T lymphocyte attenuator regulates T cell activation through interaction with herpesvirus entry mediator. *Nat Immunol,* 6**,** 90-8.

SHUI, J. W., STEINBERG, M. W. & KRONENBERG, M. 2011. Regulation of inflammation, autoimmunity, and infection immunity by HVEM-BTLA signaling. *J Leukoc Biol,* 89**,** 517-23.

SMITH, J. P., BURTON, G. F., TEW, J. G. & SZAKAL, A. K. 1998. Tingible body macrophages in regulation of germinal center reactions. *Dev Immunol,* 6**,** 285-94.

SMITH, M. R. 2003. Rituximab (monoclonal anti-CD20 antibody): mechanisms of action and resistance. *Oncogene,* 22**,** 7359-68.

SNEERINGER, C. J., SCOTT, M. P., KUNTZ, K. W., KNUTSON, S. K., POLLOCK, R. M., RICHON, V. M. & COPELAND, R. A. 2010. Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypertrimethylation of lysine 27 on histone H3 (H3K27) in human B-cell lymphomas. *Proc Natl Acad Sci U S A,* 107**,** 20980-5.

STACK, E. C., WANG, C., ROMAN, K. A. & HOYT, C. C. 2014. Multiplexed immunohistochemistry, imaging, and quantitation: a review, with an assessment of Tyramide signal amplification, multispectral imaging and multiplex analysis. *Methods,* 70**,** 46-58.

STADANLICK, J. E., KAILEH, M., KARNELL, F. G., SCHOLZ, J. L., MILLER, J. P., QUINN, W. J., 3RD, BREZSKI, R. J., TREML, L. S., JORDAN, K. A., MONROE, J. G., SEN, R. & CANCRO, M. P. 2008. Tonic B cell antigen receptor signals supply an NF-kappaB substrate for prosurvival BLyS signaling. *Nat Immunol,* 9**,** 1379-87.

STRATTON, M. R., CAMPBELL, P. J. & FUTREAL, P. A. 2009. The cancer genome. *Nature,* 458**,** 719-24.

STROUT, M. P. 2015. Sugar-coated signaling in follicular lymphoma. *Blood,* 126**,** 1871-2.

TAKATA, K., SATO, Y., NAKAMURA, N., KIKUTI, Y. Y., ICHIMURA, K., TANAKA, T., MORITO, T., TAMURA, M., OKA, T., KONDO, E., OKADA, H., TARI, A. & YOSHINO, T. 2009. Duodenal and nodal follicular lymphomas are distinct: the former lacks activation-induced cytidine deaminase and follicular dendritic cells despite ongoing somatic hypermutations. *Mod Pathol,* 22**,** 940-9.

TAKATA, K., TANINO, M., ENNISHI, D., TARI, A., SATO, Y., OKADA, H., MAEDA, Y., GOTO, N., ARAKI, H., HARADA, M., ANDO, M., IWAMURO, M., TANIMOTO, M., YAMAMOTO, K., GASCOYNE, R. D. & YOSHINO, T. 2014. Duodenal follicular lymphoma: comprehensive gene expression analysis with insights into pathogenesis. *Cancer Sci,* 105**,** 608-15.

TAN, D., HORNING, S. J., HOPPE, R. T., LEVY, R., ROSENBERG, S. A., SIGAL, B. M., WARNKE, R. A., NATKUNAM, Y., HAN, S. S., YUEN, A., PLEVRITIS, S. K. & ADVANI, R. H. 2013. Improvements in observed and relative survival in follicular grade 1-2 lymphoma during 4 decades: the Stanford University experience. *Blood,* 122**,** 981-7.

TAN, K. L., SCOTT, D. W., HONG, F., KAHL, B. S., FISHER, R. I., BARTLETT, N. L., ADVANI, R. H., BUCKSTEIN, R., RIMSZA, L. M., CONNORS, J. M., STEIDL, C., GORDON, L. I., HORNING, S. J. & GASCOYNE, R. D. 2012. Tumor-associated macrophages predict inferior outcomes in classic Hodgkin lymphoma: a correlative study from the E2496 Intergroup trial. *Blood,* 120**,** 3280-7.

UBELHART, R., BACH, M. P., ESCHBACH, C., WOSSNING, T., RETH, M. & JUMAA, H. 2010. N-linked glycosylation selectively regulates autonomous precursor BCR function. *Nat Immunol,* 11**,** 759-65.

UDUMAN, M., YAARI, G., HERSHBERG, U., STERN, J. A., SHLOMCHIK, M. J. & KLEINSTEIN, S. H. 2011. Detecting selection in immunoglobulin sequences. *Nucleic Acids Res,* 39**,** W499-504.

VAANDRAGER, J. W., SCHUURING, E., KLUIN-NELEMANS, H. C., DYER, M. J., RAAP, A. K. & KLUIN, P. M. 1998. DNA fiber fluorescence in situ hybridization analysis of immunoglobulin class switching in B-cell neoplasia: aberrant CH gene rearrangements in follicle center-cell lymphoma. *Blood,* 92**,** 2871-8.

VAN DONGEN, J. J., LANGERAK, A. W., BRUGGEMANN, M., EVANS, P. A., HUMMEL, M., LAVENDER, F. L., DELABESSE, E., DAVI, F., SCHUURING, E., GARCIA-SANZ, R., VAN KRIEKEN, J. H., DROESE, J., GONZALEZ, D., BASTARD, C., WHITE, H. E., SPAARGAREN, M., GONZALEZ, M., PARREIRA, A., SMITH, J. L., MORGAN, G. J., KNEBA, M. & MACINTYRE, E. A. 2003. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia,* 17**,** 2257-317.

VAN OERS, M. H., KLASA, R., MARCUS, R. E., WOLF, M., KIMBY, E., GASCOYNE, R. D., JACK, A., VAN'T VEER, M., VRANOVSKY, A., HOLTE, H., VAN GLABBEKE, M., TEODOROVIC, I., ROZEWICZ, C. & HAGENBEEK, A. 2006. Rituximab maintenance improves clinical outcome of relapsed/resistant follicular non-Hodgkin lymphoma in patients both with and without rituximab during induction: results of a prospective randomized phase 3 intergroup trial. *Blood,* 108**,** 3295-301.

VARGHESE, F., BUKHARI, A. B., MALHOTRA, R. & DE, A. 2014. IHC Profiler: an open source plugin for the quantitative evaluation and automated scoring of immunohistochemistry images of human tissue samples. *PLoS One,* 9**,** e96801.

VARKI, A., CUMMINGS, R. D., ESKO, J. D., FREEZE, H. H., STANLEY, P., BERTOZZI, C. R., HART, G. W. & ETZLER, M. E. 1993 (eds.) *Essentials of Glycobiology.* Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.

VARKI, A. 1993. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology,* 3**,** 97-130.

VENDEL, A. C., CALEMINE-FENAUX, J., IZRAEL-TOMASEVIC, A., CHAUHAN, V., ARNOTT, D. & EATON, D. L. 2009. B and T lymphocyte attenuator regulates B cell receptor signaling by targeting Syk and BLNK. *J Immunol,* 182**,** 1509-17.

VIARDOT, A., BARTH, T. F., MOLLER, P., DOHNER, H. & BENTZ, M. 2003. Cytogenetic evolution of follicular lymphoma. *Semin Cancer Biol,* 13**,** 183-90.

WAGNER-JOHNSTON, N. D., LINK, B. K., BYRTEK, M., DAWSON, K. L., HAINSWORTH, J., FLOWERS, C. R., FRIEDBERG, J. W. & BARTLETT, N. L. 2015. Outcomes of transformed follicular lymphoma in the modern era: a report from the National LymphoCare Study (NLCS). *Blood,* 126**,** 851-7.

WATANABE, N., GAVRIELI, M., SEDY, J. R., YANG, J., FALLARINO, F., LOFTIN, S. K., HURCHLA, M. A., ZIMMERMAN, N., SIM, J., ZANG, X., MURPHY, T. L., RUSSELL, J. H., ALLISON, J. P. & MURPHY, K. M. 2003. BTLA is a lymphocyte inhibitory receptor with similarities to CTLA-4 and PD-1. *Nat Immunol,* 4**,** 670-9.

WEIGERT, O., KOPP, N., LANE, A. A., YODA, A., DAHLBERG, S. E., NEUBERG, D., BAHAR, A. Y., CHAPUY, B., KUTOK, J. L., LONGTINE, J. A., KUO, F. C., HALEY, T., SALOIS, M., SULLIVAN, T. J., FISHER, D. C., FOX, E. A., RODIG, S. J., ANTIN, J. H. & WEINSTOCK, D. M. 2012. Molecular ontogeny of donor-derived follicular lymphomas occurring after hematopoietic cell transplantation. *Cancer Discov,* 2**,** 47-55.

WEIGERT, O. & WEINSTOCK, D. M. 2017. The promises and challenges of using gene mutations for patient stratification in follicular lymphoma. *Blood,* 130**,** 1491-1498.

WINBERG, C. D., NATHWANI, B. N., BEARMAN, R. M. & RAPPAPORT, H. 1981. Follicular (nodular) lymphoma during the first two decades of life: a clinicopathologic study of 12 patients. *Cancer,* 48**,** 2223-35.

WU, Y., EL SHIKH, M. E., EL SAYED, R. M., BEST, A. M., SZAKAL, A. K. & TEW, J. G. 2009. IL-6 produced by immune complex-activated follicular dendritic cells promotes germinal center reactions, IgG responses and somatic hypermutation. *Int Immunol,* 21**,** 745-56.

YAARI, G., UDUMAN, M. & KLEINSTEIN, S. H. 2012. Quantifying selection in high-throughput Immunoglobulin sequencing data sets. *Nucleic Acids Res,* 40**,** e134.

YAMAMOTO, T., YAMANASHI, Y. & TOYOSHIMA, K. 1993. Association of Src-family kinase Lyn with B-cell antigen receptor. *Immunol Rev,* 132**,** 187-206.

YANG, Z. Z., NOVAK, A. J., STENSON, M. J., WITZIG, T. E. & ANSELL, S. M. 2006a. Intratumoral CD4+CD25+ regulatory T-cell-mediated suppression of infiltrating CD4+ T cells in B-cell non-Hodgkin lymphoma. *Blood,* 107**,** 3639-46.

YANG, Z. Z., NOVAK, A. J., ZIESMER, S. C., WITZIG, T. E. & ANSELL, S. M. 2006b. Attenuation of CD8(+) T-cell function by CD4(+)CD25(+) regulatory T cells in B-cell non-Hodgkin's lymphoma. *Cancer Res,* 66**,** 10145-52.

YANO, T., JAFFE, E. S., LONGO, D. L. & RAFFELD, M. 1992. MYC rearrangements in histologically progressed follicular lymphomas. *Blood,* 80**,** 758-67.

YAO, S., HART, D. J. & AN, Y. 2016. Recent advances in universal TA cloning methods for use in function studies. *Protein Eng Des Sel*.

ZABALEGUI, N., DE CERIO, A. L., INOGES, S., RODRIGUEZ-CALVILLO, M., PEREZ-CALVO, J., HERNANDEZ, M., GARCIA-FONCILLAS, J., MARTIN-ALGARRA, S., ROCHA, E. & BENDANDI, M. 2004. Acquired potential N-glycosylation sites within the tumor-specific immunoglobulin heavy chains of B-cell malignancies. *Haematologica,* 89**,** 541-6.

ZHANG, J., DOMINGUEZ-SOLA, D., HUSSEIN, S., LEE, J. E., HOLMES, A. B., BANSAL, M., VLASEVSKA, S., MO, T., TANG, H., BASSO, K., GE, K., DALLA-FAVERA, R. &

PASQUALUCCI, L. 2015. Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis. *Nat Med,* 21**,** 1190-8.

ZHANG, J., VLASEVSKA, S., WELLS, V. A., NATARAJ, S., HOLMES, A. B., DUVAL, R., MEYER, S. N., MO, T., BASSO, K., BRINDLE, P. K., HUSSEIN, S., DALLA-FAVERA, R. & PASQUALUCCI, L. 2017. The CREBBP Acetyltransferase Is a Haploinsufficient Tumor Suppressor in B-cell Lymphoma. *Cancer Discov,* 7**,** 322-337.

ZHU, D., MCCARTHY, H., OTTENSMEIER, C. H., JOHNSON, P., HAMBLIN, T. J. & STEVENSON, F. K. 2002. Acquisition of potential N-glycosylation sites in the immunoglobulin variable region by somatic mutation is a distinctive feature of follicular lymphoma. *Blood,* 99**,** 2562-8.

ZOU, W. 2005. Immunosuppressive networks in the tumour environment and their therapeutic relevance. *Nat Rev Cancer,* 5**,** 263-74.

ZUCKERMAN, N. S., MCCANN, K. J., OTTENSMEIER, C. H., BARAK, M., SHAHAF, G., EDELMAN, H., DUNN-WALTERS, D., ABRAHAM, R. S., STEVENSON, F. K. & MEHR, R. 2010. Ig gene diversification and selection in follicular lymphoma, diffuse large B cell lymphoma and primary central nervous system lymphoma revealed by lineage tree and mutation analyses. *Int Immunol,* 22**,** 875-87.